A Capabilities-Based Framework for Designing and Evaluating Oil Spill Response Exercises

Authors:

Dr. Erica Gralla
Brandon Greenberg
Paule Voevodsky

with
Dr. John Harrald
Dr. Gregory Shaw
Samuel Babbitt

Final Report

February 2016

Principal Investigator: Erica L. Gralla, Ph.D.

Institute for Crisis, Disaster, and Risk Management
Department of Engineering Management and Systems Engineering
The George Washington University
800 22nd St. NW, Washington, DC 20052
Email: egralla@gwu.edu; Phone: (202) 994-3330

This page intentionally left blank.

Executive Summary

Overview

The responder community must be ready to respond quickly and effectively in the event of a marine oil spill. In order to maintain readiness, responders are required to exercise their response to oil spills on a regular basis. The critical challenge is to design, conduct, and evaluate exercises in a manner that effectively tests responders' readiness and generates lessons that can improve readiness.

This research project developed a framework to enable assessment of response readiness through evaluation of critical capabilities in exercises. The framework (1) identifies critical *capabilities* that lead to readiness for spill response, and maps them to (2) *exercise design components* that test each capability and (3) *evaluation measures* to evaluate each capability within an exercise.

The framework enables continuous improvement by linking the evaluation of exercises to the critical capabilities required of an oil spill response organization. More specifically, each exercise "tests" an organization's processes and plans. By evaluating the performance of specific capabilities, areas for improvement are clearly identified. Improvements can be made to the organization's processes and plans and then tested in the next exercise. In this manner, continuous improvement is enabled and evaluated, and response readiness is improved.

Framework

Element 1: Capabilities *are activities or skills required during an oil spill response.* Successful performance of capabilities requires the *capacities* to marshal appropriate resources, the *competencies* required to manage the effort, and the *abilities* to perform required activities or skills (Harrald, 2006).

The set of capabilities includes *functional capabilities,* which are operational activities that responders must conduct during the response (such as containing the spill or notifying stakeholders), *management and support capabilities* which support and coordinate the performance of functional capabilities (such as coordinating operations or managing resources), and *skills-based capabilities*, which are individual and team-oriented capabilities that enable responders to work together effectively and to respond flexibly to the situation as it develops (such as managing teams or solving problems).

Identifying a set of key capabilities provides a focus for improvement and evaluation efforts, including the design and evaluation of exercises. Each capability is defined, then relevant exercise design components and evaluation measures are provided.

Element 2: Exercise Design Components *are events, issues, problems, or requirements incorporated into an exercise, which require and challenge the performance of capabilities and enable them to be evaluated.*

The set of exercise design components includes aspects of the *baseline event* such as the precipitating spill, its location, and the weather; *baseline tasks* such as defining goals or holding a meeting; and *complexity factors* that challenge participants in five ways: *dealing with the unexpected, dealing with scale and time, managing influences and conflict, organizing and managing people,* and *dealing with ambiguity*.

Each exercise design component is described and mapped to the capability or capabilities it tests. Exercise planners can integrate several of these exercise design components into an exercise scenario designed to test particular capabilities.

Element 3: Evaluation Measures are the criteria (what to evaluate), metrics, and techniques (how to evaluate) that enable the evaluation of the performance of a capability within an exercise.

Two or more evaluation criteria are provided for each capability; they provide clear guidelines for evaluating the performance of each capability. Exercise designers can prepare a checklist-style evaluation sheet for exercise observers to fill out in order to evaluate the exercise.

Contributions

The key contribution of the framework described in this report is that the framework:

- Enables assessment of response readiness through exercises, and in particular, enables a cycle of continuous improvement, by linking exercise design and evaluation to specific response capabilities.

Additionally, the framework:

- Expands the set of capabilities evaluated in exercises. The framework includes the key capabilities identified in NPREP, but also adds more management and support capabilities and additional skills-based capabilities that enable organizations to respond to unimagined and unplanned-for events.

- Provides a toolbox for consistent and comprehensive exercise design. Using the framework ensures consistency and comparability across exercises, even with different or new exercise designers.

- Enables exercises to be scored on specific dimensions of performance (capabilities) rather than passing or failing the entire exercise, enabling clear and actionable identification of areas requiring improvement.

- Enables evaluation of an operator's or the entire industry's improvement (or lack thereof) over time, and holds response organizations accountable for improvement, by re-testing capabilities in a series of exercises.

Table of Contents

1 Introduction

1.1 Project Overview

1.1.1   Motivation: Exercises that build and assess readiness

The responder community must be ready to respond quickly and effectively in the event of a marine oil spill. However, given the rarity of spills, responders have few opportunities to practice skills, refine processes, and learn lessons through experience. In order to maintain readiness, therefore, responders are required to exercise their response to oil spills on a regular basis. Oil spill response exercises have progressed significantly in the past several decades, from practicing skills and tactics to exercising coordination and management.

The next step is to ensure that exercises build and assess readiness for response. Readiness is a complicated concept, which must be evaluated across an interacting system of organizations, resources, and technologies, at various organizational levels. In this research, we focus on understanding how to assess readiness in specific types of exercises, to build a framework that will be useful more broadly.

The critical challenge is to design, conduct, and evaluate exercises in a manner that effectively tests responders' readiness and generates lessons that can improve readiness. Current exercise guidelines do not suggest specific exercise design elements and evaluation measures that improve readiness ("Draft NPREP Guidelines," 2015, "NPREP Guidelines," 2002). While important lessons have been learned through decades of exercise experience, most focus on high-level planning rather than on the details of each exercise scenario and how it is evaluated (e.g., Cashman, 2011). There is a need for systematically cataloging and/or developing exercise design components and evaluation techniques that are specifically designed to test readiness and generate implementable lessons that will lead to continual improvements in oil spill response readiness.

1.1.2   Project deliverable: A capabilities-based framework for designing and evaluating exercises

To address this challenge, this research has developed a framework that (1) identifies critical *capabilities* that lead to readiness for spill response, and maps them to (2) *exercise design components* that test each capability and (3) *evaluation measures* to evaluate each capability within an exercise.

Element 1: Capabilities *are activities or skills required during an oil spill response.* Successful performance of capabilities requires the *capacities* to marshal appropriate resources, the *competencies* required to manage the effort, and the *abilities* to perform required activities or skills (Harrald, 2006).

1

The set of capabilities includes *functional capabilities,* which are operational activities that responders must conduct during the response (such as containing the spill or notifying stakeholders), *management and support capabilities* which support and coordinate the performance of functional capabilities (such as coordinating operations or managing resources), and *skills-based capabilities*, which are individual and team-oriented capabilities that enable responders to work together effectively and to respond flexibly to the situation as it develops (such as managing teams or solving problems).

Identifying a set of key capabilities provides a focus for improvement and evaluation efforts, including the design and evaluation of exercises. Most evaluation frameworks focus solely on abilities and somewhat on capacity; our research has developed an exercise framework that integrates all three concerns (we label them all "capabilities" for the sake of simplicity).

Element 2: Exercise Design Components *are events, issues, problems, or requirements incorporated into an exercise, which require and challenge the performance of capabilities and enable them to be evaluated.*

The set of exercise design components include aspects of the *baseline event* such as the precipitating spill, its location, and the weather; *baseline tasks* such as defining goals or holding a meeting; and *complexity factors* that challenge participants in five ways: *dealing with the unexpected, dealing with scale and time, managing influences and conflict, organizing and managing people,* and *dealing with ambiguity*.

Each exercise design concept is described and mapped to the capability or capabilities it tests. Exercise planners can integrate several of these exercise design components into an exercise scenario designed to test particular capabilities.

Element 3: Evaluation Measures *are the criteria (what to evaluate), metrics, and techniques (how to evaluate) that enable the evaluation of the performance of a capability within an exercise.*

Two or more evaluation criteria are provided for each capability. They provide clear guidelines for evaluating the performance of each capability. Each evaluation criterion suggests one or more measurement techniques (methods for capturing the relevant data), such as having an observer rate the performance, discussing it during a post-exercise hotwash, or asking participants to rate themselves.

Exercise designers can prepare a checklist-style evaluation sheet for exercise observers to fill out in
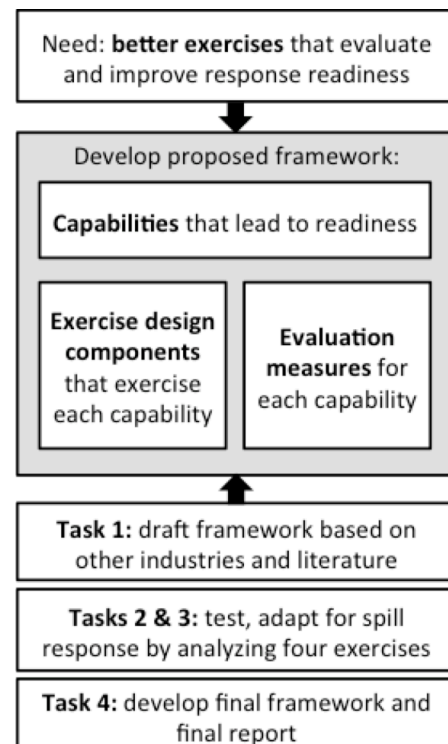
Figure 1: Summary of research project

order to evaluate the exercise.

Linkages Among Framework Elements: Each *capability* is linked to (1) several *exercise design components* which prompt the performance of that capability, and (2) two or more *evaluation criteria* for that capability (what to measure), along with suggested *evaluation techniques* (how to measure). Conversely, each *exercise design component* is linked to one or more *capabilities* that it prompts during an exercise, and each *evaluation technique* is linked to one or more *capabilities*. A matrix shows all the links among the framework elements.

Updating the Framework: The framework is intended to be a living document, in that BSEE (or others) should, as needed, update definitions, change links, add elements, or make other modifications as the document is used and as our collective understanding of exercise design and evaluation – and of the key capabilities for spill response – evolves over time.

1.1.3   Research approach

The framework was initially drafted based on our experience with exercises in various industries and on a survey of the literature. Then, it was adapted specifically for oil spill response by testing and refining it through observation of four oil spill response exercises. Figure 1 summarizes the research approach. Section 2 of this report explains the research methods in more detail.

1.1.4   Using the framework

Figure 2 illustrates the envisioned usage of the framework. An exercise designer would begin by selecting a set of capabilities to practice in an exercise from those listed in the framework. Next, the exercise designer would look up the set of exercise design components mapped to each capability, and select one or more to include in the exercise. In addition, the exercise designer would look up the set of appropriate evaluation measures, and select one or more. The exercise would be designed and run, then evaluated using the evaluation techniques to measure performance on each tested capability.
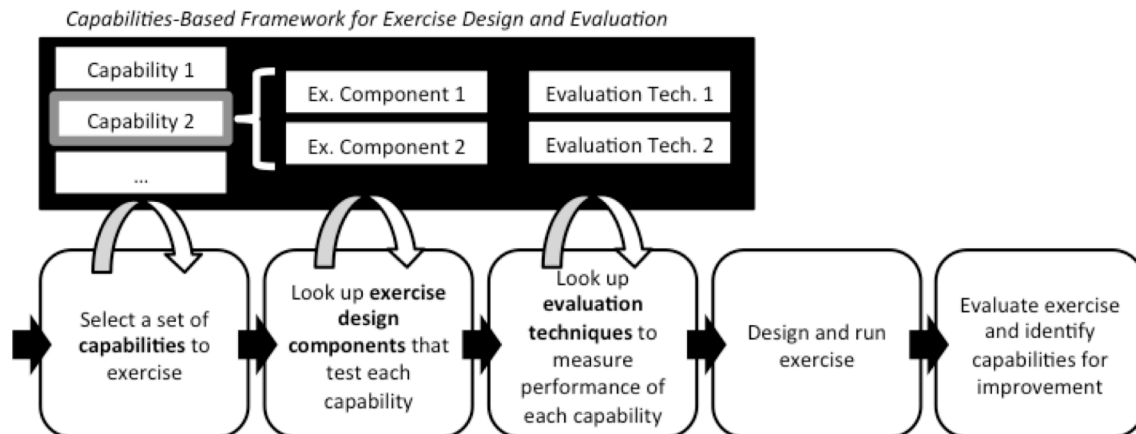
Figure 2: Envisioned usage of framework

3

If used, the framework will provide the tools to ensure that each exercise is effective in testing and improving oil spill response readiness. Future guidelines could require the testing of various sets of capabilities within the suite of required exercises.

An illustrative example of using the framework is provided later in this report, in Section 4.2.

1.2 Project Contributions

1.2.1 Enabling continuous improvement

The critical challenge that is addressed by this project is to design, conduct, and evaluate exercises in a manner that effectively tests responders' readiness and generates lessons that can improve readiness.

Our framework enables continuous improvement by linking the evaluation of exercises to the critical capabilities required of an oil spill response organization.

Figure 3 illustrates a notional improvement cycle. An organization has some response processes and plans in place. An exercise "tests" these processes and plans, enabling some evaluation of how well they performed. Based on these evaluations, the organization can identify areas of poor or mediocre performance and thereby identify areas for improvement. These improvements can be made to the organization's processes and plans, and tested in the next exercise. In this manner, continuous improvement is enabled and response readiness is improved.

To enable this cycle, our framework defines critical *capabilities* for oil spill response, *exercise design components* that test those capabilities and *evaluation measures* that reveal performance on those capabilities.
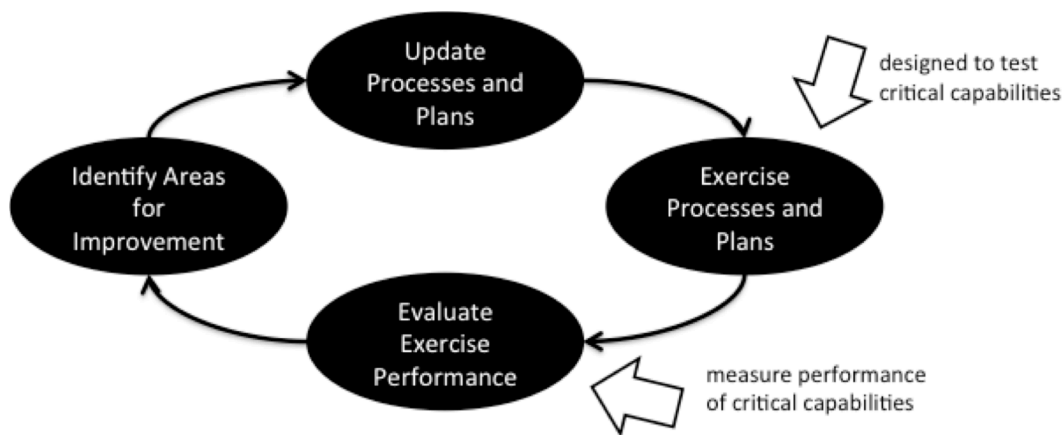
Figure 3: Continuous improvement in spill response readiness

1.2.2   Envisioned contributions of the framework

The key contribution of the framework described in this report is that the framework:

- Enables assessment of response readiness through exercises, and in particular, enables the cycle of continuous improvement described in Figure 3, by linking exercise design and evaluation to specific response capabilities.

Additionally, the framework:

- Expands the set of capabilities evaluated in exercises. The framework includes the key capabilities identified in NPREP, but also adds more management and support capabilities that are key to response success and additional skills-based capabilities that enable organizations to respond to unimagined and unplanned-for events.

- Provides a toolbox for consistent and comprehensive exercise design. Using the framework ensures consistency and comparability across exercises, even with different or new exercise designers.

- Enables exercises to be scored on specific dimensions of performance (capabilities) rather than passing or failing the entire exercise, enabling clear and actionable identification of areas requiring improvement.

- Enables evaluation of an operator's or the entire industry's improvement (or lack thereof) over time, and holds response organizations accountable for improvement, by re-testing capabilities in a series of exercises.

Finally, the framework lays a foundation for accomplishing the following:

- Defining "success" in performing a capability. The capability definitions and evaluation measures are a starting point for describing precisely what it means to execute capabilities successfully.

- Evaluating readiness for response to unexpected elements in the next disaster. By evaluating capabilities related to flexibility and adaptability, and by adding complexity to existing scenarios, the framework enables evaluating not only whether an organization can accomplish a specific capability but also how well it can do so, on how large a scale it can do so, and whether it can do so in more and more complex conditions.

- Updating processes and regulations to require performance of clearly defined capabilities within exercises, to clearly defined evaluation standards.

1.3 Related Work

This section describes previous work that is related to this research project and identifies the gap in existing knowledge that this project fills.

It is difficult to evaluate whether exercises demonstrate response readiness, in part because the concept of "readiness" has not been clearly defined. Capabilities-based approaches to exercise design and evaluation can avoid some of these difficulties. If capabilities that enable readiness are defined, then these capabilities can be evaluated within an exercise. Capabilities-based frameworks work well for four reasons. First, organizational response structures are increasingly organized around specific capabilities to help manage the response ("NRF," 2013). Second, they help evaluators understand specific performance issues that may affect readiness (U.S. Department of Homeland Security, 2013). Third, they enable capabilities to be linked to specific performance objectives, similar to instructional design methodologies (Gagné, Wager, Golas, & Keller, 2005). Finally, they enable capability-based objectives beyond performance to be included in an exercise. For example, these might include objectives related to training, education, planning, and/or relationship-building (Jackson & McKay, 2011).

Existing research studies and exercise programs have outlined two different types of capabilities relevant for both disaster and oil spill response. The first type of capabilities are soft-skill and managerial competencies and abilities, such as leadership and development, coordination, decision making and communications. Response personnel and organizations need these capabilities in order to response to different disaster types and sizes (Biddinger et al., 2008; Brady, 2003; Dausey, Buehler, & Lurie, 2007). The second type of capabilities are functional (e.g., technical or task-oriented) capabilities that often need to be performed during a specific type of response ("Draft NPREP Guidelines," 2015, "NPREP Guidelines," 2002, "U.S. Coast Guard Incident Management Handbook," 2001). These include capabilities such as public information and warning, shoreline cleanup, or spill containment.

However, existing research and best practices do not provide clear guidance on how to design exercises so that capabilities are performed nor on how to ensure evaluation is strongly linked and attributed to capability performance. This linkage is critical because the design of an exercise can impact responders' decisions and actions, which could then impact the performance of the different capabilities being evaluated (Berlin & Carlström, 2015). For example, an exercise that does not require much coordination could result in poor or excellent performance of the coordination capability due to, respectively, a lack of realistic coordination challenges or an overly simple scenario.

Existing research and programs have focused on several areas that do not specifically or strongly link capabilities to exercise design and evaluation. First, much literature has focused on processes and methodologies to evaluate performance (without linkages to specific capabilities) (Beamon & Balcik, 2008; Cooper et al., 2010; Gebbie, Valas, Merrill, & Morse, 2006; Rosen et al., 2010). Second, there has been significant discussion on the merits of different exercise types (e.g., simulation, tabletop, operations-

based, etc.) but not on design elements that test the performance of capabilities (Chi, Chao, Chuang, Tsai, & Tsai, 2001; Dausey et al., 2007; Descatha et al., 2009; Rosen et al., 2010; Zhou et al., 2015). Third, there has been a focus on how exercises achieve learning outcomes but not performance outcomes (Arnold et al., 2009; Fautua, Schatz, Reitz, & Bockelman, 2014; Silenas, Akins, Parrish, & Edwards, 2008). Lastly, there has been a focus evaluating the execution of plans rather than overall performance ability (Cashman, Stephens, & Boyles, 2003; Franks, Knutson, Parker, & LeJeune, 2011).

While there has been some focus on identifying the "right" set of capabilities for response ("NRF," 2013, "TCL," 2007), there has been little focus on how to cohesively prompt and then evaluate the performance of these capabilities within exercises. Evaluators may try to fill this gap based on their own experience, which can introduce variability in exercise design and evaluation that makes it difficult to generalize or compare evaluations. Achieving a systematic and consistent evaluation of capabilities requires the development of a way to prompt capabilities within exercises and evaluate their performance. This research is intended to fill this gap.

1.4 Remarks: Design Philosophy of this Project

In this section, we explain several philosophies that guided the development of this project, and explain its relationship to other similar programs and goals.

1.4.1   Focus on evaluation and continuous improvement

This research focuses on evaluation rather than on training. Training emphasizes improving individual and organizational knowledge, skills and abilities through the course of a single exercise. Exercises designed to train might be constructed differently from those designed to evaluate. For example, training focuses on instilling and practicing the most commonly needed skills, whereas an exercise with an evaluative goal might test reactions to unexpected challenges. Training and participant learning is not a focus of this research project.

However, the ultimate goal is to enable *learning*, or continuous improvement, over the course of many exercises. By evaluating exercises, areas for improvement will be identified. The idea is that each operator must then train or otherwise improve their ability, which in turn will be evaluated in the next exercise. Thus, we focus on enabling learning over time rather than on meeting specific learning objectives within an exercise.

1.4.2   Focus on evaluating exercises through capabilities

Evaluation is a complicated topic in research and practice. Evaluating the success of an exercise can be accomplished in many different ways and with many different goals.

The focus of this research has been on developing methods to prompt and evaluate specific capabilities that need to be performed during an oil spill response. Good demonstrations of these capabilities allow exercise evaluators to understand the strengths and weakness of exercise participants. The goal is to enable an evaluation of the

capabilities of the operators, rather than the overall success of the outcome of an exercise. We chose to focus on this goal, i.e., on evaluating operator capabilities, because it enables the larger goal of evaluating operator readiness for oil spill response. It is possible that an operator could complete a successful exercise and still be unable to complete a successful real response, but if an operator has successfully demonstrated an important suite of capabilities, it is more likely that they will be able to successfully deploy these capabilities in reality.

However, this framework is only a first step in understanding how to evaluate the success of an entire exercise and, more importantly, understanding whether and how this relates to the likelihood of success in a real oil spill response.

1.4.3 Relationship between performance in an exercise and in a real response

This research focuses on identifying specific capabilities that are needed in an oil spill response and developing measures that indicate how well they were performed *in an exercise*. The implicit assumption is that capabilities that are successfully performed in an exercise would also be performed successfully in a real response, which in turn would lead to a successful response outcome. This assumption is the basis for many exercise programs as well as this research. However, this assumption has not been explicitly validated (or invalidated) in the literature or in practice. Our focus was on capabilities that are likely needed for a real-world response and evaluating how well they were performed in a simulated environment.

As a result, the measures developed in this research do not measure outcomes that would be achieved in a real-world response. This is because many response outcomes cannot be measured based on an exercise due to its artificiality, scope, and timeframe. For example, we do not measure loss of life or the economic impact of a spill (key outcomes indicating the success of a real response) because these data are not available within the artificial environment of the exercise. There are two related and important questions. First, what makes an exercise successful? Second, what does a successful exercise indicate about the likelihood of successful response outcomes? These important questions about response "readiness" are not addressed in this research. Rather than focusing on outcomes, we focus on capabilities, for two reasons: first, they can be measured in an exercise, and second, it is reasonable to assume that capabilities demonstrated in an exercise would also be demonstrated in a real response.

1.4.4 Relationship to other exercise design frameworks

We are not the first researchers to develop the idea of a capabilities-based framework for exercise design; indeed, the philosophy of designing exercises around capabilities has been practiced extensively for many years. Our contribution is in designing a framework that enables this type of exercise design, and in adapting it for oil spill response specifically.

The framework developed in this research project follows the philosophy of instructional design (Gagné et al., 2005), which is also the basis for other government exercise programs, such as the Homeland Security Exercise and Evaluation Program (U.S. Department of Homeland Security, 2013). These programs also recommend capabilities-based design of exercises and evaluation frameworks. The ADDIE model is the most common instructional design model. It begins by identifying the desired objectives and capabilities, then building the scenario around them. Evaluations of the group, individual, and exercise should be performed to identify gaps in both training and policy to improve preparedness prior to an actual event. Subsequent assessments should be utilized to determine the quality of the exercise as a learning tool and identify weaknesses in the training scenario. This process is also recommended by FEMA for the National Exercise Program.

Our framework uses the same philosophy: that exercises should be designed to test specific capabilities. However, it does not envision such an in-depth process of crafting exercises for more than one use, but rather focuses on the design of individual exercises to test capabilities, and on how these individual exercises build up into an exercise program that enables continuous improvement (see Section 4.3).

1.4.5 Implementability of the framework

The framework described in this document is intended to be useful to BSEE and to be implementable immediately without significant additional resources (see Section 4.3 for more details on this topic). However, we were not constrained by the current regulatory environment. We included some capabilities and exercise design components that might be problematic for BSEE to evaluate under current regulations. There is no obligation for BSEE to evaluate or utilize these capabilities and components. We felt it was important to include them because they are relevant to increasing readiness for oil spill response. It is possible that they will be useful to other organizations or to BSEE in the future.

1.5 Structure of this Document

This document describes the results of the research project "Developing a Capabilities-Based Framework for Designing and Evaluating Oil Spill Response Exercises." It begins with an introduction (Section 1), which includes the project overview (Section 1.1), contributions (Section 1.2), related work (Section 1.3), and remarks on the design philosophy of the project (Section 1.4). Next, the methods used to develop the framework are described in Section 2.

The framework itself is summarized in Section 3, including the capabilities (Section 3.1), exercise design components (Section 3.2), evaluation measures (Section 3.3), and the linkages among them (Section 3.4). In this summary, the definition of each type of element is given, the scope and design philosophy are described, and the organization of each set of elements is explained. The section includes a list of the elements included in the framework, but the detailed definition for each capability, each exercise design component, and each measure are provided in the appendices.

Section 4 explains how to use the framework, including the steps required (Section 4.1), an illustrative example (Section 4.2), requirements for implementing the framework (Section 4.3), and remarks on designing exercise programs using the framework (Section 4.4).

The document closes with a set of recommendations (Section 5) based on our experience during the research project and several ideas for future work (Section 6) that build on the framework and/or our experience carrying out the project.

The appendices include the detailed definitions of all of the capabilities (Appendix A), exercise design components (Appendix B), and evaluation techniques (Appendix C). Each capability includes a definition, a list of relevant evaluation measures, and a list of relevant exercise design components. To link exercise design components and evaluation techniques back to the relevant capabilities, a matrix is provided in Section 3.4.

2 Methods: Developing the Framework

Two major steps were employed in developing the framework. First, a draft framework was developed based on the literature. Second, it was adapted for oil spill response by observing and analyzing four oil spill response exercises.

2.1 Developing a Draft Framework Based on the Literature

The first step was to develop a draft version of the framework based on the extant literature. We surveyed literature in related industries, such as international humanitarian response and domestic emergency response. In surveying the literature, we sought papers and reports that discussed learning from simulation exercises, performance measurement, and critical skills for response.

Based on our reading of the literature, we aimed to identify capabilities required for response, exercise design components, and exercise evaluation techniques. While our framework aims to link these to one another, the literature does not link them, so we identified independent sets of each. We first identified large numbers of each, then distilled the large list into a shorter one by combining similar items and refining the descriptions.

Next, we examined the sets of capabilities, exercise design components, and evaluation techniques. We identified and added missing elements to each set by drawing on our experience (including that of our senior advisors) and by attempting to map capabilities to design components that test each capability. Finally, we organized the elements into categories or types.

Finally, we held additional discussions with our senior advisors, who provided feedback on the descriptions of the framework elements.

2.2 Refining the Framework and Adapting it for Oil Spill Response

The second major task in the research project was to refine the draft framework and adapt it specifically for oil spill response. We did so in several steps, described below.

The performance of an oil spill response exercise cannot be evaluated by examining only the outcome; the *process* of achieving it must also be investigated in order to evaluate the degree of success in critical capabilities and identify areas for improvement. We utilized

social science methods that allowed us to "see" and record the process of the response exercise and identify the main factors leading to the outcomes (Langley, 1999). These qualitative methods are particularly useful when processes and outcomes depend on factors such as management skills that cannot be easily quantified.

The method involves recording actions in detail, "coding" the data to identify critical themes or actions, then analyzing these data to identify the main factors that led to outcomes. These methods are further explained in the paragraphs that follow. These methods have previously been successfully utilized to study related industries, such as wildland firefighting (Weick, 1993) and responses to problems in nuclear and chemical plants (Carroll, Rudolph, & Hatakenaka, 2002).

2.2.1   Observing oil spill response exercises

A first step was to observe four oil spill response exercises.

We selected exercises based on three criteria. First, we sought exercises that exemplify the core of BSEE's exercise program. Second, we selected exercises based on their timing and the ability of the research team to observe them. Third, we sought variety in several dimensions, including the spill management team contractor, the exercise timeline (such as whether it focused on the first few hours or days 3-4), the scale of the exercise (varying from small tabletop exercises to large area exercises), the elements in play (such as source control or equipment deployment), and whether the exercise was operator-led, government-led, and/or unannounced.

We observed four exercises:
1. Operator-led SMT tabletop exercise, beginning day 1 of a major spill with source control in play.
2. Government-Initiated Unannounced Exercise (GIUE) with SMT tabletop and equipment deployment components.
3. Government-Initiated Unannounced Exercise (GIUE) with SMT tabletop and source control components.
4. Full-scale Area Exercise, beginning a few days after a major spill with shoreline impact imminent.

In each exercise, the research team (three researchers in three of the four exercises; two researchers in the fourth) attended and observed each exercise and recorded hand-written notes. The researchers focused on collecting data that described the process of the response, including key events, actions, decisions, information gathered and shared, and outcomes achieved. Observers sought to record the key discussion points and decisions related to the oil spill response. They identified information as it flowed among people and organizations, and noted the information on which decisions were based. In addition, they tracked the progress of the response through key events and outcomes. The result was a set of notes from each researcher describing their observations. Each researcher observed different parts of the exercise (for example, in one case, one researcher observed the source control efforts while the other observed the spill management team).

The notes from each researcher were consolidated into one chronology of the spill response for analysis.

2.2.2　Data analysis techniques

To analyze the data, we employed qualitative coding techniques. Qualitative data analysis techniques call for "coding" data incidents so that they can be categorized, counted, and related to one another (Langley, 1999; Miles & Huberman, 1984). For this purpose, we used the Atlas.ti software that enables qualitative data to be stored, coded, categorized, and analyzed.

2.2.3　Data analysis: Refining the set of capabilities

The first step in the data analysis process was to refine the set of capabilities from the draft framework based on the data from the observed exercises.

The list of capabilities from the draft framework was used to "code" the data by identifying incidents in the data in which the exercise participants were demonstrating a capability. For example, we might apply a code for "transfer knowledge from past experience" to an incident in which a responder suggested a tactic that had worked previously but modified it for the current application. The same code might be applied to another incident in which a responder shared information with someone outside her organization because she recalled that in a past response, someone with that role needed the information. In this manner, we identified the variety of incidents in which a given capability might be needed and demonstrated.

Additional codes were created to capture any capabilities that appeared to contribute to successful performance but were not yet included in the draft framework; thus, the set of capabilities was expanded as needed.

The definitions of each capability were refined based on the observations from the exercise, to more precisely define what the capability entailed and what made it successful.

2.2.4　Data analysis: Refining the set of exercise design components

The second step in the data analysis process was to refine the set of exercise design components from the draft framework based on the data from the observed exercises.

The notes from each exercise were analyzed to identify different events, issues, problems or requirements that prompted participants to perform capabilities, or that challenged the exercise participants. Definitions for each of the resulting exercise design components were developed and refined as the analysis progressed. In some cases, exercise design components were combined or separated to better reflect the differences and similarities between them.

Next, we linked exercise design components to specific capabilities. From our earlier data analysis, we had identified ("coded") incidents in the notes that described participants

executing a capability. We re-analyzed the notes that were coded for each capability. If an event or situation coded for a capability was created by a particular exercise design component, the relevant exercise design component was "assigned" to the capability. Through this process, each capability accrued a set of relevant design components.

We were also able to record the number of times each component was found prompting the capability. Based on these frequencies and on more detailed understanding of how each exercise design component prompted each capability within the observed exercises, we sorted the exercise design components for each capability by their relevance (strong, moderate, and weak relevance). Definitions for each of the exercise design components were again developed, refined, combined and/or separated as the data were analyzed.

2.2.5   Data analysis: Creating evaluation measures

The third step is to identify *evaluation measures* to measure the performance of each capability within the exercise. To do so, we again started with each coded incident that demonstrates a particular capability, then developed two or more criteria that can evaluate the coded capability's performance based on data from the exercise. By this point, based on the preceding analysis, we had a large number of coded data incidents that demonstrate each capability, from all four exercises. By examining all of them together, we were able to identify the various kinds of incidents that demonstrate the capability. Then, we built evaluation criteria that can be measured based on what we saw in the exercises.

2.2.6   Feedback from senior advisors

Lastly, the project's senior advisors reviewed the definition of each element of the framework to ensure they were clear, concise, and appropriate to the oil spill response context. They also reviewed the complete sets of elements to identify missing elements. Finally, they reviewed the linkages between the elements of the framework. The framework was adjusted based on their feedback.

3 Framework Summary

The following sections summarize each of the three sets of elements of the framework: capabilities, exercise design components, and evaluation measures. For each one, we provide a definition, describe the scope and design philosophy that guided the design of the set of elements, and summarize the set of elements included in the framework. The full description of each element is provided in the appendices.

3.1 Capabilities

3.1.1   Definition of a capability

A capability is an activity or skill required during an oil spill response. Successful performance of a capability requires the *capacities* to marshal appropriate resources, the *competencies* required to manage the effort, and the *abilities* to perform required activities or skills (Harrald, 2006). Capabilities may include straightforward activities or tasks as well as more complex management and skills-based abilities.

3.1.2   Scope and design philosophy

We sought to capture not only the basic *abilities* to perform required activities or skills, but also the *capacities* to marshal all the required resources and the *competencies* required to manage the entire effort. By broadening our focus in this manner, we were able to identify a more comprehensive set of capabilities.

Identifying all the critical capabilities for response readiness is a major undertaking, because there are myriad skills, processes, and competences that may be required in response to widely varying spill scenarios. To manage the scope of this research project, we focused on three areas:
1. Basic capabilities for spill response laid out in NPREP
2. Capabilities related to management and support
3. Capabilities that enable organizations to be effective and flexible in dynamic and challenging environments.

Regarding (1), basic capabilities for spill response laid out in NPREP, our framework captures those relevant to BSEE's mission. A comparison between the capabilities required by NPREP and those laid out in our framework is provided below, in Section 3.1.4.

Regarding (2), capabilities related to management and support, our framework describes the activities that enable the basic NPREP capabilities to be executed, by providing support and coordination.

Regarding (3), capabilities that enable organizations to be effective and flexible in dynamic and challenging environments, our framework captures capabilities inspired by

the literature on crisis response. Many of these capabilities are "intangible," in that they are not straightforward activities, but rather processes or skills that must be utilized throughout the course of an exercise. Examples include problem solving skills, adaptability, and improvisation.

The NPREP capabilities were included because they are already identified as critical for oil spill response and are core to BSEE's mission.

Our framework also goes beyond NPREP in two important areas. We focused on management and support capabilities because they are equally critical but less well explained in NPREP and other documents. Finally, we added capabilities that enable flexibility and effectiveness in dynamic environments because these types of skills have been shown to be critical when responding to unpredictable disasters, in a variety of industries.

3.1.3   Summary and organization of capabilities in the framework

The response capabilities in this document provide a comprehensive set of capacities, competencies and abilities that need to be performed by responders during an oil spill emergency. Response capabilites are organized into three main categories:

1. **Functional capabilities** are specific activities that responders must conduct in order to sucessfully respond to an oil spill emergency.

2. **Management and support capabilities** are activities that support and coordinate the performance of functional capabilities.

3. **Skills-based capabilities** are individual and team-oriented capabilities that enable responders to work together effectively and to respond flexibly to the situation as it develops.

The functional capabilities are needed when responding to an oil spill and are typically evaluated based on whether and/or how well they were executed. However, the management and support and the skills-based capabilities are essential in enabling organizations to plan for and execute functional capabilities in large, complex, and uncertain working environments that are typical in many oil spill responses.

The list below summarizes the final set of response capabilities identifed in the research. Related sub-capabilites were also identified and are included in the following table. Descriptions of each capability are provided in Appendix A.

Functional Capabilities
- Deploy and Demonstrate Use of Equipment
- Ensure Safety and Security
    - Execute Emergency Procedures
- Inform and Warn the Public
- Notify and Update Stakeholders and Response Partners

- Contain Spill to Mitigate Environmental Impact
- Mitigate Economic Impact of Spill
- Control and Stop Oil at Source
- Fight Fires in a Marine Environment
- Conduct Human Search and Rescue
- Conduct Wildlife Search and Rescue
- Conduct Shoreline Cleanup and Restoration
- Conduct Salvage and Lightering
- Separate and Dispose of Oil and Debris

Management and Support Capabilities
- Set and Monitor Progress toward Goals and Objectives
- Conduct Strategy and Response Planning
- Assess and Monitor the Evolving Situation
  - Assess Technical Hazards
  - Project Future Issues and Concerns
- Utilize Management Structure with Operating Norms
  - Assign/Delegate Responsibilities
  - Scale Operations
  - Maintain and Manage Compliance with Laws and Regulations
- Interface with and Manage Stakeholders
- Coordinate Operations
- Provide and Manage Operational Communications
- Manage and Share Information
  - Retain Data and Information
  - Manage Continuous Updates
  - Create and Maintain Shared Situation Awareness
  - Utilize Public Information and Intelligence
- Manage and Account for Resources
  - Source and Procure Resources
  - Mobilize Response Resources
  - Track and Report on Personnel, Assets, and Finances
  - Maintain and Service Equipment
  - Support Response Personnel

Skills-Based Capabilities
- Deploy and Manage Effective Teams
- Solve Problems and Make Decisions
  - Evaluate Alternatives
- Adapt and Be Flexible
- Utilize Adaptive Expertise
- Improvise
- Prioritize Response Efforts
- Utilize Prior Plans to Support Operations

3.1.4 Comparison of capabilities with NPREP

Table 1 shows the capabilities identified in the existing NPREP guidelines and the corresponding capabilities developed from the research.

The framework captures all the NPREP capabilities, and also adds many capabilities that are not present in NPREP. NPREP focuses primarily on functional capabilities, with a few management and support capabilities. We include many more management and support capabilities, and we add several skill-based capabilities related to readiness for unexpected and unplanned-for events.

Table 1: Comparison between NPREP and our framework

NPREP Capabilities	Framework Capabilities
Coordinate a Response	Coordinate Operations
Demonstrate Usage of Equipment	Deploy and Demonstrate Use of Equipment
Notify Stakeholders	Notify Response Partners and Stakeholders
Conduct Salvage and Marine Firefighting	Conduct Salvage and Lightering; Fight Fires in a Marine Environment
Execute Emergency Procedures	Execute Emergency Procedures
Remote Assessment and Consultation for Vessels for Salvage and Marine Firefighting	Assess and Monitor the Evolving Situation; Fight Fires in a Marine Environment
Organizational Design Capabilities	
Notifications	Notify Response Partners and Stakeholders
Staff Mobilization	Manage and Account for Resources (Mobilize Response Resources)
Ability to Operate within Response Management System in Plan	Utilize Management Structure with Operating Norms
Operational Response Capabilities	
Discharge Prevention/Control	Control and Stop Oil Spill at Source
Assessment of Discharge/Vessel Stabilization/Fire Suppression	Assess and Monitor the Evolving Situation; Fight Fires in a Marine Environment; Conduct Salvage and Lightering
Containment of Discharge/Vessel Stabilization/Fire Suppression	Contain Spill to Mitigate Environmental Impact; Fight Fires in a Marine Environment; Conduct Salvage and Lightering
Recovery of Spilled Material	Conduct Shoreline Cleanup and Restoration; Conduct Marine Cleanup and Restoration;
Protection of Sensitive Areas	Conduct Strategy and Response Planning; Contain Spill to Mitigate Environmental Impact
Disposal of Recovered Material and Contaminated Debris	Separate and Dispose of Oil and Debris
Response Support Capabilities	
Communications	Provide Operational Communications
Transportation	Manage and Account for Resources (Mobilize Response Resources)
Personnel Support	Manage and Account for Resources (Support Response Personnel); Ensure Security and Safety
Equipment Maintenance and Support	Manage and Account for Resources (Maintain and Service Equipment)
Procurement	Manage and Account for Resources (Source and Procure Resources)
Documentation	Manage and Share Information

3.2 Exercise Design Components

3.2.1 Definition of an exercise design component

An exercise design component is an event, issue, problem, or requirement incorporated into an exercise that requires and challenges the performance of capabilities and enables them to be evaluated.

Exercise design components are a new term not well discussed in existing literature or policy documents. However, the concept is intuitively understood by exercise designers that have tried to make exercises more realistic and challenging for exercise participants.

3.2.2 Scope and design philosophy

The development of the set of exercise design components was guided by the philosophy that the set should include the basic elements of the scenario, but also enable exercise designers to add complexity and challenges to the exercise in specific ways. By adding complexity and, at the same time, assessing the performance of specific capabilities, the exercises will be better able to assess readiness for spill response in a variety of conditions and for a variety of challenging capabilities.

The goal was to create a "menu" for exercise designers to select from when putting together the exercise scenario. The unique details of each scenario are left to the designer to develop: the exercise design components do not suggest detailed ocean currents or specific unexpected events. Instead, the "menu" includes the need for a current and an unexpected event, leaving the exercise designer to write in the details. Given the many possible scenarios, such a catalog would have been unwieldy.

A second area that is out of scope is the implementation of the exercise scenario. Once the scenario is designed, the exercise designer must determine what is provided at the start of the exercise and what is injected during the exercise (and when and how).

3.2.3 Summary and organization of exercise design components in the framework

The exercise design components in this document provide guidance and suggestions for an exercise designer developing the scenario that prompts exercise participants to practice particular capabilities. A well-constructed scenario creates a more realistic operating environment and enhances learning and assessment for exercise participants.

For example, the "decision making" capability might be practiced in response to a dilemma in which participants must decide whether to prioritize the stoppage of a spill or concentrate on spill impact mitigation because of limited resources. In order to mimic the difficulties of prioritizing efforts in a real response, components such as unclear situation or problem, stakeholder influence, incomplete or conflicting information, and

interpersonal conflict may be introduced to simulate conditions that decision makers will need to navigate.

A comprehensive exercise scenario incorporates three types of exercise design components that are described in the following sections. They are built around the following "model" of a prototypical exercise, which contains three types of components:
1. Baseline event components
2. Baseline task components
3. Complexity factors

An exercise consists of a *baseline event* that initiates and continues the exercise. To respond, participants must complete a series of *baseline tasks*. These first two exercise components constitute the baseline exercise scenario, which is then made more complex by adding *complexity factors* to make the exercise more challenging for participants. For example, a straightforward decision-making task can be made more difficult if the need for the decision is not clear to the participants; in response, they must demonstrate more complex capabilities, such as identifying problems to be solved. Figure 4 illustrates the relationships between the design components and the exercise scenario.
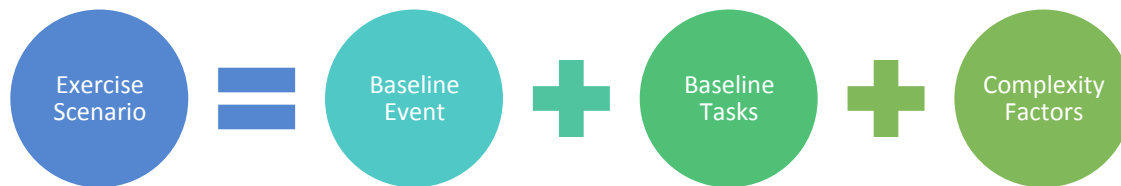
Figure 4: Three types of exercise design components make up an exercise scenario

The following sections discuss the three different sets of exercise design components that prompt participants to demonstrate capabilities.

Exercise Design Components: Baseline Event
Baseline event exercise design components comprise a comprehensive description of the emergency event and the general conditions in which the response occurs. The description 1) provides critical information to exercise participants, and 2) aligns exercise designers and evaluators about the "ground truth" of the exercise. Three exercise components are used to describe the baseline event:
1. Specify Precipitating Event
2. Specify Event Impacts and Locations
3. Specify Environmental Conditions

Exercise designers can make the event more complex or challenging by incorporating additional complexity into their event descriptions. For example, by incorporating plan or strategy conflicts, exercise participants will have to go through the additional step of de-conflicting or developing new plans or strategies, which can take significant time and effort. Event descriptions should therefore be used in conjunction with the complexity factors described below.

Exercise Design Components: Baseline Tasks
Tasks are activities that exercise participants must accomplish. Baseline tasks are activities that are typical or common in a real-world response. Completion of these tasks are not by themselves indicators of a successful response, but it would be difficult to achieve a successful response without them. They stimulate participants to complete various sets of actions and generally included in operational plans or regulations. In addition, many of these tasks are already incorporated, in some form, by the command structure and processes (e.g., the Incident Command System) used to respond. Additional tasks may be designed into the exercise.

Many tasks are relatively straightforward, but can be made more complex depending on the conditions in which the task must be accomplished (see the next category of exercise design components). For example, developing a plan of action may be straightforward when one person does it alone with complete information. However, it is more difficult to develop a plan in stressful conditions, such as doing so with a geographically dispersed team within two hours based on incomplete or conflicting information. Baseline tasks should therefore be used in conjunction with the complexity factors described below.

The baseline tasks described in Appendix B are generic. For example, one baseline task is "Require Plans or Strategies." Our framework does not describe the various types of plans or strategies that may be required; these details are left to the exercise designer. Additionally, we only include baseline tasks that were identified in research. There may be additional tasks not identified in the research that may be relevant to oil spill response exercises.

Exercise Design Components: Complexity Factors
Complexity factors are exercise design components that are incorporated into the baseline event or baseline tasks to increase the complexity or difficulty of the exercise scenario. They are particularly useful for exercise participants who have mastered capabilities in relatively straight-forward and uncomplicated response environments, but need additional challenges to ensure they can respond in more realistic and complex environments. These exercise design components are grouped into five main categories:
1. Dealing with the Unexpected
2. Dealing with Scale and Time
3. Managing Influences and Conflict
4. Organizing and Managing People
5. Dealing with Ambiguity

One or more of the complexity factors may be utilized within an exercise scenario. For example, an exercise designer may choose to incorporate several different types of complexity factors (e.g., resource insufficiency, change in situation, and ambiguous "correct" decision") and/or incorporate several of the same complexity factors (e.g., multiples types of resource insufficiencies).

The list below summarizes the final set of exercise design components identifed from the research. Complexity factors challenge exercise participants in five areas, which are also

shown in the table. Descriptions of each exercise design component are provided in Appendix B.

Baseline Event
- Specify Precipitating Event
- Specify Event Location and Impact
- Specify Environmental Conditions

Baseline Tasks
- Require Plans or Strategies
- Require Execution of Plans, Procedures or Processes
- Require Group Decision Making
- Require Meetings
- Requie Updates, Reports, or Information
- Require Strategic, Operational or Tactical Decisions
- Require Authorizations
- Require Deployment and Demonstration of Equipment

Complexity Factors
- Dealing with the Unexpected
  - Incorporate an Unannounced Exercise or Withheld Scenario
  - Incorporate Resource Insufficiencies
  - Incorporate Unexpected Situations
  - Incorporate Changes in the Situation
- Dealing with Scale and Time
  - Incorporate Events or Tasks that Vary in Size and Require Different Response Levels
  - Incorporate Multiple Events
  - Incorporate Time Flow Adjustments
  - Incorporate Time Pressure
- Managing Influences and Conflict
  - Incorporate Stakeholder Influence on Priorities, Goals or Tasks
  - Incorporate Public Perception of Priorities, Goals or Tasks
  - Incorporate Conflicting Organizational Goals and Objectives
  - Incorporate Events or Tasks that Require Prioritization of Response Efforts
  - Incorporate Plan or Strategy Conflicts
- Organizating and Managing People
  - Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel
  - Incorporate Ambiguous Group Structures and Operating Norms
  - Incorporate Interpersonal Conflicts
  - Incorporate Team-Based Decisions or Actions
- Dealing with Ambiguity
  - Incorporate Unclear Situations or Problems
  - Incorporate Incomplete or Conflicting Information

- Incorporate Ambiguous "Correct" Decisions
- Incorporate Events or Tasks that Stress Plans
- Incorporate Events or Tasks that Exceed Experience or Expertise of Responders

3.3 Evaluation Techniques and Measures

3.3.1 Definitions of evaluation techniques and measures

Evaluation measures are the criteria (what to evaluate), metrics, and techniques (how to evaluate) that enable evaluation of the performance of a capability. These elements are used concurrently to evaluate the performance of each capability.

More specifically, the criteria are specific questions defined for each capability that enable the evaluation of that specific capability in an exercise. Each criterion is accompanied by simple metrics. Evaluation techniques are methods for capturing the data required to use the criteria. The most common technique is for an exercise observer to rate the participants' performance of the capability, but several other techniques are possible, such as discussing performance during a hotwash.

3.3.2 Scope and design philosophy

Capability performance can be measured any number of ways. We considered measures of input, process, output, and outcome, to ensure a comprehensive evaluation of performance of each capability.

On the other hand, utilizing several performance criteria for each capability would quickly become unwieldy when evaluating an exercise for several capabilities. We included approximately two criteria for each capability, though in a few cases three are suggested. The aim was to include one measure of the process responders followed and one measure that was associated with the outcome of interest. Outcomes can be more difficult to assess because they are not always easily observable. For example, assessing whether a decision's outcome was successful is not straightforward, since it requires a definition of success and perhaps an understanding of what would have happened if a different decision had been made. It is more straightforward to assess the process – whether participants made decisions by considering several alternatives and utilizing all available information – because the criteria for evaluation are clearer. In other cases, even when the process itself cannot be easily observed, the outcome of that process may be visible. A second reason for including a process measure and an outcome measure was that it enables evaluators to note when correct processes were followed, even if the outcome was not perfect (perhaps due to other problems).

The list of measurement techniques was drawn from a survey of the literature on measuring performance within exercises. All relevant techniques identified from the survey were included.

3.3.3 Summary and organization of evaluation measures and techniques in the framework

Exercises are a performance-based training and assessment tool. Exercises enable capability demonstration without going through a real-world scenario, and they enable assessment of the preparedness of a complex response organization. To do so effectively, the measures and approaches used to evaluate exercises must be appropriate for measuring response readiness. One important goal is to assess the likelihood that the demonstrated capabilities can be performed again in a variety of conditions. This is a difficult task that requires a thoughtful approach to evaluation.

As described earlier, there are two parts to the measures. First, evaluation criteria and associated metrics are defined for each capability; they enable the evaluation of that specific capability in an exercise. Evaluation *techniques* are methods for capturing the data required to use these measures.

Evaluation Measures (Criteria and Metrics)
Evaluation criteria and metrics are provided for each of the capabilities included in the framework. Two to three measures are included for each, as described above. Please refer to Appendix A, the set of capabilities, to see the measures developed for each capability.

Evaluation Techniques
This section describes the different techniques that can be used to collect and analyze data to evaluate the performance of capabilities in exercises. These techniques can be used separately or together depending on the goals of the exercise. Each technique has strengths and weaknesses that may be better suited for evaluating different capabilities. Additionally, data for the purposes of exercises include any observations, documents, discussions, or notes that can be used to understand the how well a capability was demonstrated.

The list below summarizes the final set of evaluation techniques identified in the research. Descriptions of each evaluation technique are provided in Appendix C.

Evaluation Techniques
- Observer rating
- Probing/Question Asking
- Proxies and Indicators
- Self-Rating
- Post-Hoc Modeling
- Evaluation of Plans
- Documentation
- Hotwash and Debriefing
- After Action Review

3.4 Linkages Among Framework Elements

As described earlier, an important aspect of the framework is the ability to link capabilities to exercise design components that prompt the performance of the capability and to evaluation measures that enable evaluation of that performance. The linkages among these elements are summarized in the matrix below. In addition, each capability listed in Appendix A is explicitly linked to exercise design components and to evaluation measures.

The linkages from capabilities to exercise design components and evaluation techniques are categorized by their degree of relevance to the capability. The three categories of relevance are defined as:

- **S = Strong.** The exercise design component directly stimulates or stresses ability to perform capability. Evaluation techniques are well suited to evaluating capability.
- **M = Moderate.** The exercise design component can stimulate or stress ability to perform capability, but not as the primary prompt for the capability. Evaluation techniques can be used to evaluate capability, but not as primary evaluation method.
- **W = Weak.** Design component can tangentially stimulate or stress ability to perform capability. Evaluation techniques can be used to evaluate the capability, but not as a primary evaluation method.

The linkages among elements are summarized in three figures, for clarity. Each figure maps the capabilities to either exercise design components or evaluation techniques. Figure 5 maps the capabilities to relevant baseline event and baseline task exercise design components. Figure 6 maps the capabilities to relevant complexity factor exercise design components. Figure 7 shows the capabilities mapped to relevant evaluation techniques.

Capabilities	Baseline Event			Baseline Tasks							
	Specify Precipitating Event	Specify Event Location and Impact	Specify Environmental Conditions	Require Plans or Strategies	Require Execution of Plans, Procedures or Processes	Require Group Decision Making	Require Meetings	Require Updates, Reports or Information	Require Strategic, Operational or Tactical Decisions	Require Authorizations	Require Deployment and Demonstration of Equipment
Functional	S	S	S	S	S	M	M	M	S	S	S
Management and Support											
Set and Monitor Progress toward Goals and Objectives	S	S	M	S			S	S	S	W	
Conduct Strategy and Response Planning	S	S	S	S	M	S	M	S	S	S	W
Assess and Monitor the Evolving Situation	M	S	S	M	W	M	M	S	S	M	
Utilize Management Structure with Operating Norms	S	S	M	S	M	M	S	S	S	S	W
Interface with and Manage Stakeholders	S	S	W	S	S	S	S	S	S	S	
Coordinate Operations	M	S	S	W	W				S	M	
Provide and Manage Operational Communications		M	S	M	M			W	M		S
Manage and Share Information	M	S		S	S	S	S	S	S	S	
Manage and Account for Resources	S	S	S	M	W		M	M	S	M	S
Skills-Based											
Deploy and Manage Effective Teams	S	S	W	M		W	S	W	W	M	
Solve Problems and Make Decisions	S	S	S	S	W	S	M	S	S	S	W
Adapt and Be Flexible	M	S	S	W					W		W
Utilize Adaptive Expertise	M	S	S	W					M		W
Improvise	W	M	S								
Prioritize Response Efforts	M	S	S	M		S	M	M	S	M	W
Utilize Prior Plans to Support Operations	S	S	S	S	W	W	W	W	S	M	

Figure 5: Capabilities mapped to baseline event and baseline tasks exercise design components

Table: Capabilities mapped to complexity factor exercise design components. Cell values: S (Strong), M (Moderate), W (Weak); blank = no mapping.

Complexity Factors	|	Skills-Based →							Management and Support →								
		Utilize Prior Plans to Support Operations	Prioritize Response Efforts	Improvise	Utilize Adaptive Expertise	Adapt and Be Flexible	Solve Problems and Make Decisions	Deploy and Manage Effective Teams	Manage and Account for Resources	Manage and Share Information	Provide and Manage Operational Communications	Coordinate Operations	Interface with and Manage Stakeholders	Utilize Management Structure with Operating Norms	Assess and Monitor the Evolving Situation	Conduct Strategy and Response Planning	Set and Monitor Progress toward Goals and Objectives
Dealing with the Unexpected																	
Incorporate Resource Insufficiencies	S	M	S	S	M	S	S	W	S	M	M	S	M	M	W	S	S
Incorporate Unexpected Situations	S	M	W	M	M		S	M	S	S	W	M	M	M	M	S	S
Incorporate Changes in the Situation	S	M	M	M	W	S	S	M	S	S	W	M	S	S	W		S
Incorporate an Unannounced Exercise or Withheld Scenario	S	S	S		S	S	M	M	S	W	W		S				S
Dealing with Scale and Time																	
Incorporate Events or Tasks that Vary in Size and Require Different Response Levels	M	M	S			M		S	S	S	M	S			S	M	
Incorporate Multiple Events	S		S				M	W	M	M	W			S	M	M	
Incorporate Time Flow Adjustments								W		W							
Incorporate Time Pressure	M			M						W				W	W	W	
Managing Influences and Conflict																	
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	W	S	S		M	S	S	W	W	S	W	M	S	M	S	W	W
Incorporate Public Perception of Priorities, Goals or Tasks	W	W	S			M	S	W	M	M	W		S	W	M	M	W
Incorporate Conflicting Organizational Goals and Objectives		W	S				S	W				M		S			M
Incorporate Events or Tasks that Require Prioritization of Response Efforts	M	S	S			S	S	S	S				S	S	S		
Incorporate Plan or Strategy Conflicts	M	M	M	M	W	S	W	M	W			M		S	M	M	
Organization and Managing People																	
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse	S	S	S		S	S	S	S	S	M		S	S	S	W		
Incorporate Ambiguous Group Structures and Operating Norms	M				M		M	S	M			M	S	M	W	W	
Incorporate Interpersonal Conflicts	W					M	S	W	W			W		W	S	W	
Incorporate Team-Based Decisions or Actions	W			M		S	S	S	W			M		W	M	S	
Dealing with Ambiguity																	
Incorporate Unclear Situations or Problems	S		M		S	W		M	M	M	W		M		S		
Incorporate Incomplete or Conflicting Information	M	S	M	M	S	M	S	M	S	S	W	S	S	S		M	
Incorporate Ambiguous "Correct" Decisions	M	M	S	M	W	M		M	W	S	M	S		M		M	
Incorporate Events or Tasks that Stress Plans	S	S		M	S	M	S	M	M			W	W	W		S	
Incorporate Events or Tasks that Exceed Experience or Expertise of Responders	S	W	W	M	M	S	S	W						W		S	

Figure 6: Capabilities mapped to complexity factor exercise design components

Capabilities	Observer Ratings	Probing/Question Asking	Proxies & Indicators	Self-Rating	Post-Hoc Modeling	Evaluation of Plans	Documentation	Hotwash and Debriefing	After Action Review
						Evaluation Techniques			
Functional	S		S		M	M	M	S	S
Management and Support									
Set and Monitor Progress toward Goals and Objectives	S	S	M	W				M	M
Conduct Strategy and Response Planning	S			W			M	M	S
Assess and Monitor the Evolving Situation	S	S		M			S	M	S
Utilize Management Structure with Operating Norms	S	S	S	M	M		S	M	S
Interface with and Manage Stakeholders	S	S	M	W			S	M	S
Coordinate Operations	S	S		M	M			M	S
Provide and Manage Operational Communications	S		M					M	S
Manage and Share Information	S	S		W			M	M	S
Manage and Account for Resources	S	S	M	W			M	M	S
Skills-Based									
Deploy and Manage Effective Teams	S		M	M				M	S
Solve Problems and Make Decisions	S	S	M	M				M	S
Adapt and Be Flexible	S	S		M				W	S
Utilize Adaptive Expertise	S	S		M				W	S
Improvise	S	S		M				W	S
Prioritize Response Efforts	S	S		M			M	W	S
Utilize Prior Plans to Support Operations	S	S							S

Figure 7: Capabilities mapped to relevant evaluation techniques

4 Using the Framework

4.1 Using the Framework to Design an Exercise

To use the framework to design an exercise, four basic steps must be completed. They are summarized here and explained in more detail in the illustrative example, below.

1. **Select the capabilities to be tested in the exercise**

   Several capabilities should be selected for each exercise. Each exercise could form part of a series, so that each exercise re-tests capabilities previously found to be weak, and tests increasingly complex capabilities over time.

2. **Look up and select exercise design components that test these capabilities**
   a. **Include baseline event components**
   b. **Include baseline tasks (optional)**
   c. **Include complexity factors (optional)**

   For each capability, relevant exercise design components should be looked up and a tentative list of selected components should be created. The list must include baseline event components, and may optionally include baseline tasks and complexity factors.

3. **Create a scenario and specific tasks for the exercise, utilizing these exercise design components**
   a. **Create the baseline event scenario**
   b. **Add baseline tasks and complexity factors to the scenario**
   c. **Decide how to implement the scenario (what is injected and when)**
   d. **Determine the need for involvement from other stakeholders**

   The exercise designer should write a detailed scenario inspired by the selected design components and designed to test the selected capabilities. The scenario must include the baseline event components, and may include baseline tasks and complexity factors. The final step is to decide how to implement the scenario, including what information is provided to participants at kickoff and what is injected later, when, and how. If any involvement is required from other stakeholders, that should be arranged as well.

4. **Compile measures for evaluating capability execution in the exercise**
   a. **Review and compile measures to evaluate each capability**
   b. **Select evaluation techniques for each measure**
   c. **Decide pass/fail criteria**

For each capability, relevant evaluation measures should be looked up and selected. The measures can be copied to an evaluation sheet for evaluators to rate during or after the exercise. In addition, pass/fail criteria for the exercise should be determined; for example, in order to pass, what grade is required on each tested capability?

4.2 Illustrative Example

The illustrative example below explains all of the steps to designing an exercise, which are summarized above.

4.2.1   Step 1: Selecting the capabilities to be tested

When designing an exercise, the first step is to choose the capabilities on which the responders will be evaluated. These will be the bones around which the exercise will be designed.

For the purpose of this example, we will choose two capabilities:
- The Skill-Based Capability "Prioritize Response Efforts", and
- The Functional Capability "Deploy and Demonstrate Use of Equipment."

The selection of the capabilities to be tested could either be done on a case by case basis by an individual exercise designer, or it could be part of a series of exercises designed to build and evaluate specific capabilities for each operator or for the industry as a whole. Section 4.4, below, suggests ways to build up an exercise program.

4.2.2   Step 2: Selecting exercise design components that test these capabilities

Looking up exercise design components for each capability
When the capabilities have been chosen, the next step is to look up the exercise design components that test these capabilities. There are two ways to do this: they can be looked up via the Linkage Matrix, or through each selected capability. The Linkage Matrix relates all capabilities to the complete list of design components with which the capabilities can be tested. Alternatively, the definition of each capability is followed by a list of the exercise design components which test that capability.

In either case, the linked list of exercise design components includes baseline event components, baseline tasks, and complexity factors. The linkages between the capability and the exercise design components are categorized as strong, moderate, and weak.

At this point, a set of exercise design components can be tentatively selected (the set can be modified later as the scenario details are written, in Step 3).

Including baseline event components

At minimum, the baseline event components – which includes the precipitating event, the event impact and location, and the environmental conditions – must be included in the exercise. The Linkage Matrix shows that the capability "Prioritize Response Efforts" is strongly tied to "Event Impacts and Environmental Conditions," and that "Deploy and Demonstrate Use of Equipment," as a Functional Capability, is strongly connected to all three baseline event components. Therefore, these exercise design components must be included in the scenario to test these capabilities.

Including baseline tasks

In addition, the exercise scenario could require additional baseline tasks. Many tasks are already included in the exercise by default, because they are required by the Incident Command System processes, but further baseline tasks could be tentatively added to the list of selected components. In our example, we select the baseline task "Require Strategic, Operational, or Tactical Decisions" in order to test the capability "Prioritize Response Efforts."

Including complexity factors

Finally, the exercise designer might select additional complexity factors to stress the responders to demonstrate each capability or to make its execution more challenging. For example, to prompt the responders to demonstrate their capability "Prioritize Response Efforts," the designer might select the exercise design component "Incorporate Multiple Events," which has a strong tie to the capability. To add further challenges related to the same capability, the designer might also include "Incorporate Events or Tasks that Require Prioritization of Response Efforts," and "Incorporate Resource Insufficiencies."

At this point, a tentative list of exercise design components has been selected, including the baseline event components and several complexity factors.

4.2.3   Step 3: Create a scenario and specific tasks for the exercise

The exercise designer's next task is to create a detailed scenario for the exercise, utilizing the selected exercise design components. The exercise design components suggest an outline for the exercise, but it is up to the exercise designer to create the details of the events that take place in the scenario.

Exercise designers have by now selected the tasks or actions they want performed; next, they must decide how they want to implement the requirement for those tasks or actions within the exercise. For example, injecting a request that US Fish & Wildlife would like to know how you plan to clean up contaminated wildlife would create the need for developing a wildlife cleanup plan. Designers can be explicit about what they want to see or it can be implicitly contained in an inject.

Creating the baseline event scenario

The first step is to write a scenario around the baseline events. All the baseline event components must be included in each exercise: the precipitating event, the event location and impact, and the environmental conditions. For instance, to test the capabilities

selected for our example, the designer might come up with a scenario including a spill at location [X], near the shore, at 0800 hours. The spill has released [A] gallons of oil, and is continuing to release oil at a rate of [B] gallons per hour. The spilled oil is moving in the [NW] direction at [C] miles per hour.

Adding baseline tasks and complexity factors to the scenario
From here, the designer can add baseline tasks and complexity factors to the scenario in order to strain the responders to display more desired capabilities.

To increase the complexity of the scenario and prompt the responders to display their capability to "Prioritize Response Efforts," the designer might consider the selected exercise design components "Incorporate Events or Tasks that Require Prioritization of Response Efforts" and "Incorporate Resource Insufficiencies." A scenario inspired by these components is that the oil spill is large enough so that the boom available to the responders cannot feasibly cover all areas in need of protection. As a result, the exercise participants will have to demonstrate the "Prioritize Response Efforts" capability.

Considering the selected baseline task component "Require Strategic, Operational, or Tactical Decisions," the exercise designer might write into the scenario that after an hour, the movement of the spill has changed. It had initially been moving toward a lower-priority area but is now headed toward a high-priority location. This scenario utilizes not only the selected baseline task but also the complexity factor "Incorporate Changes in the Situation." (Note that while this complexity factor was not initially selected, it is nevertheless relevant. While the scenario is being written, it is worth re-examining the Linkage Matrix for additional relevant exercise design factors.)

The capability "Deploy and Demonstrate Use of Equipment" is a little more self-explanatory in the relevant design components related to its observation. The designer would most likely literally "Require Deployment and Demonstration of Equipment" in order to test whether the responders can adequately display this capability, but might also add components to stress the scenario, such as leaving the decision of which type of equipment to deploy up to the responders ("Require Strategic, Operational and Tactical Decisions") or placing the spill relatively close to priority areas so the responders must deploy the equipment quickly and/or navigate through shallow waters ("Incorporate Time Pressure", "Require Drill and Practice of Particular Skills").

Deciding how to implement the scenario
It is also important that the designer decide when to inform responders of the aspects of the scenario, meaning what will be described in the kickoff and what will be injected throughout the exercise. Kickoff documents, injections, and other relevant documents must be written.

Depending on exercise objectives, exercise designers may choose provide varying levels of detail to exercise participants to facilitate play from the start of an exercise. For example, exercise designers may choose to release additional information about the event throughout the exercise in order to better mimic real-world disasters in which all

information about the event is not known initially. However, a comprehensive event description is still needed to ensure exercise design and execution efforts are aligned.

For example, the designer might include the baseline event descriptions in the kickoff along with the size of the spill and the amount of boom available, but inject the change in spill trajectory an hour into the exercise.

Determining the need for involvement from other stakeholders
Additionally, it is important that designers determine the need for involvement from other agencies or stakeholders. For instance, to further test the "Prioritize Response Efforts" capability, the designer might require collaboration with a government organization such as the United States Fish and Wildlife Service, which directly involves the design components "Incorporate Events or Tasks that Require Coordination Among Dispersed & Diverse Personnel" and "Incorporate Stakeholder Influence on Priorities, Goals or Tasks," both of which have a strong connection with the capability.

4.2.4   Step 4: Compile measures for evaluating capability performance in the exercise

Reviewing and compiling measures to evaluate each capability
The final step is to compile the measures that will be used to evaluate capability performance during the exercise. The simplest way to do this is to look up each capability that is to be tested in this exercise, and review the evaluation measures listed for that capability in the framework. One or more of the available measures should be included. We recommend including at least one easy-to-observe measure and one measure associated with the outcome or output.

The measures are formatted so that they can be copied into an evaluation sheet for the exercise. The evaluation sheet should list all the measures, and evaluators can use it to rate performance.

Selecting the evaluation technique
An additional step is to select the techniques for evaluation. All the measures list suggested measurement techniques. Virtually all of them can be measured by an observer of the exercise, but some can also be measured in other ways, such as discussion during a hotwash. For each measure, the exercise designer should determine how performance will be measured.

Examples
In our example, to evaluate the "Prioritize Response Efforts" capability, there are three measures listed. Some or all of these can be included in the evaluation sheet. We recommend including at least one easy-to-observe measure and one measure associated with the outcome. In this case, for example, an easy-to-observe measure is the "Process" measure: "recognize the need to prioritize response efforts and establish a process and timeline in which to prioritize events." During the exercise, an observer could rate its success. A measure associated with the "outcome" is the responders' "ability to prioritize response efforts [in a way that] positively impacts response operations." This can be rated

through observation during the exercise, but it can also be evaluated in a number of other ways, including a participant hotwash. A hotwash is useful here because the evaluator can directly ask the participants how they think they did in their prioritization of response efforts, and their perceptions can be valuable and informative when evaluating their performance.

In our example, to evaluate the "Deploy and Demonstrate Use of Equipment" capability, the output measure of whether the responders "deploy appropriate response equipment" should be observed during the exercise and measured on a Yes/No basis. Additionally, responders must "appropriately and safely demonstrate use of response equipment. If unable to perform, they discuss how they would operate the equipment." Evaluators should observe this during the exercise, and measure success as 1) Did Not Perform Well, 2) Performed Satisfactorily, or 3) Performed Well.

Determining pass/fail criteria
The final step is to set the criteria for an organization to pass the exercise. We suggest that operators be required to achieve a certain minimum evaluation on each tested capability, or on a subset of the tested capabilities, in order to pass the exercise. For example, an operator might pass for a particular capability if they achieved "Performed Satisfactorily," or "Performed Well," but not if they achieved "Did Not Perform Well."

This framework is not designed to enable a single score to be calculated for the exercise as a whole based on the scores for the capabilities tested. We do not provide methods for calculating an overall score because we do not know the goals of each exercise. Exercises designed to test management capabilities should be scored on the success of several management and support capabilities, but exercises designed to test equipment deployment should be scored differently. It is up to each exercise designer to determine the requirements for "passing" an exercise based on the evaluations achieved in each of the evaluated capabilities.

4.3 Implementing the Framework: Requirements

An important question is whether any additional resources are required to utilize the framework as recommended in this report.

In exercise design, *no additional resources are needed*. The exercise design process is similar to current practices. The key changes to the process are (1) exercise goals are determined by selecting capabilities to be tested, and (2) exercise design components are used as inspiration during the writing of the exercise scenario.

In exercise evaluation, *no additional resources are needed.* The exercise evaluation process is similar to current practices. The key changes to the process are (1) measure capability performance rather than overall exercise performance, using the provided measures; (2) create, print, and use evaluation sheets for the tested capabilities; and (3) we recommend (but it is not required to) use the evaluation measures to structure the

hotwash discussion. Minimal additional training is required to use the new evaluation measures, because they are largely self-explanatory.

4.4 Designing Exercise Programs Using the Framework

4.4.1 Designing a series of exercises

The framework is most powerful when it is used over a series of exercises to enable the continuous improvement shown in Figure 3, above. When an exercise is designed to test critical capabilities and measure how well they are performed, the operator receives clear feedback on capabilities they must improve. The next exercise should re-test those capabilities to ensure that they have indeed improved.

A sample exercise program is shown in Table 2. In this example, the goals are to ensure the operator has basic abilities not only in executing a response but also in managing it, then to build skills in flexibility and adaptability to ensure the operator can respond to unexpected aspects of future spills.

As the exercise program progresses, the exercises test different capabilities but also re-test those tested previously, to ensure that the operator's capabilities are improving. The capabilities tested can increase in complexity, in order to push organizations to continuous improvement.

We suggest designing exercise programs to achieve goals for industry preparedness, such as ensuring that all operators are proficient in management and support capabilities and have some expertise in adapting to unexpected situations. A similar idea is suggested by the Homeland Security Exercise and Evaluation Program (U.S. Department of Homeland Security, 2013). It suggests designing an exercise program based on overall objectives, then designing a series of exercises that satisfies these objectives.

Table 2: Sample exercise program

Sample Exercise Program
Exercise program goals: 1. Ensure operator has basic abilities in executing a spill response   2. Ensure operator has basic abilities in managing a spill response   3. Build skills in flexibility and adaptability

Exercise 1: Basics
 • Test two functional capabilities:
   ○ Contain Spill to Mitigate Environmental Impact
   ○ Control and Stop Oil Spill at Source
 • Test one management and support capability
   ○ Assess and Monitor the Evolving Situation

Exercise 2: Reinforce basics and add management   • Re-test the same capabilities from Exercise 1

• Test two additional management and support capabilities
○ Coordinate Operations
○ Manage and Account for Resources
Exercise 3: Reinforce management and add flexibility
• Re-test the two capabilities added to Exercise 2
• Test two new skills-based capabilities
○ Adapt and Be Flexible
○ Utilize Adaptive Expertise

If exercise programs are designed consistently across many operators, it will be possible to assess the state of readiness of the industry overall. In cases like oil spills, where events happen rarely, it is possible that the state of preparedness degrades through lack of practice. Assessing a well-defined set of capabilities through a consistent program of exercises can describe the overall industry state of preparedness and how it is changing over time.

4.4.2 Recommendation: Evaluate key capabilities and share with operators

To enable continuous improvement through a series of exercises, our recommendation is to evaluate operators for their performance on the capabilities tested in the exercise, whether or not they lead to passing or failing the exercise as a whole. These evaluations can be retained by BSEE and also communicated to the operators, in order to give targeted feedback to the operators. At the next exercise for each operator, their previous evaluations for these capabilities should be compared to their new evaluations, so that it is clear whether operators are improving. In this manner, BSEE could hold operators accountable for learning from exercises and continuing to improve.

The overall exercise pass/fail decision could rest on only a subset of these capabilities. A subset of "core" capabilities could be designated (for example, encompassing only those included in the NPREP guidelines), and a poor performance on any of these could lead to an exercise fail, but poor performance on others would not cause exercise failure.

BSEE's current practice is to give operators a "pass" or "fail" evaluation for each exercise. While this practice is consistent with current regulations, it represents a missed opportunity for learning and improvement from exercises. When operators pass, they receive limited and unstructured feedback on what could be improved, and they have little incentive to invest in improvements. Operators are not asked in future to show improvement on those areas which were identified as problems. Using our framework in the manner recommended here would help to improve on this situation.

5 Recommendations

The following recommendations are based on our observations during the project, on the literature examined for this research, and on our team's experience with exercises in the areas of U.S. homeland security and international disaster response.

These recommendations focus on improving the exercise program for two main goals:
1. Evaluate an operator's ability to execute is response plan
2. Generate lessons learned to enable continuous improvement in response readiness

While BSEE's major responsibility lies with the first goal, the second goal is clearly in BSEE's interest to promote.

5.1 Recommendations: Current Exercise Program

The following recommendations are relevant to the current exercise program, in that they should not require major changes to regulations or policy.

5.1.1 Expand and guide the post-exercise hotwash

BSEE's current practice is to conduct a hotwash after each exercise. In most cases, BSEE personnel and any other invited stakeholders (such as Coast Guard personnel) spend time apart from the operator and contracted SMT to go over any problems or issues they saw. Next, they join the participants for the hotwash. Typically, the participants are first given an opportunity to note any issues that arose or points of improvement; then, the evaluators do the same.

While this practice has led to some important lessons, it is somewhat ad hoc, so it does not result in consistent and systematic learning from the exercises. Participants usually take it seriously, but often focus on one major takeaway without considering or looking for other lessons beyond the most obvious. A second problem is that the hotwash is often conducted in a hurry, since participants and evaluators expect a half-day exercise and often have other commitments later in the day.

Our recommendation to address these problems is to add some structure to the hotwash. Rather than simply asking for observations and takeaways, evaluators could use a checklist and ask for participants' and evaluators' evaluations of performance in several important dimensions. Our framework provides a number of possibilities: performance could be assessed on the critical capabilities tested in that exercise. With this structure, participants and evaluators would be prompted to think about lessons learned in several key areas, enabling broader lessons to be learned. A second recommendation is to plan for a lunch or coffee break before the hotwash, so that participants and evaluators are refreshed and ready to spend time thinking through the lessons learned.

5.1.2   Recruit stakeholders to participate in exercises

BSEE currently attempts to recruit key stakeholders to participate in exercises. For example, we observed several exercises with Coast Guard and local government participation. However, these outside stakeholders are not always able to participate, and when they do not, their actions are "simulated," and these simulated actions typically align with the SMT's goals. For example, an SMT might seek approval for dispersant use. If CG is not participating, their approval will be assumed.

Our observations suggest that the participation of outside stakeholders is critical to testing the SMT's ability to execute the response plan, and to generating important lessons. In one exercise observed by the research team, CG was not participating, so it was assumed that approval for dispersant use was granted when it was asked for. However, in another exercise, CG personnel were participating. Approval for dispersant use was not granted when requested, and additional information was sought from the spill management team in a different format. A long exchange between the SMT and CG resulted in both parties learning more about the required process of getting dispersant approval. This interaction with the stakeholders was critical to learning this important lesson.

Our recommendation is therefore that BSEE continue to recruit key stakeholders to participate in exercises, and to step up these efforts where possible. In addition, recruiting multiple stakeholders to each exercise makes them even more valuable.

5.1.3   Include equipment deployment in SMT exercises

BSEE's unannounced exercises include equipment deployment only occasionally. Of three exercises we observed, only one required any equipment deployment.

While including equipment deployment requirements is expensive for the operators, our observations suggest that it significantly enhances BSEE's ability to evaluate the operator's ability to execute the plan. When equipment need not be deployed in reality, its deployment and operation are typically assumed to proceed as planned. However, this is rarely the case in reality. Equipment may break down, or have difficulty finding its target location, or be unable to communicate, or have other problems. More importantly, managing these problems and communicating with the vessels adds significantly to the SMT's overall workload.

Our recommendation is that BSEE include equipment deployment as part of the SMT exercises rather than conducting these drills separately, to the extent possible. It is important to include this level of complexity in the scenario not only so that the SMT can be evaluated in managing real resources but also to ensure the SMT can manage this increased and more complex workload. It is much less challenging to do one exercise focused on equipment deployment and another focused on the SMT than to include the equipment deployment within the SMT exercise.

5.1.4 Include realistic stress and pressure

BSEE's current practice already includes important challenges in that exercises are often conducted unannounced, so the SMT has no preparation for the exercise. However, most exercises were conducted beginning at the start of a workday, and included breaks for evenings. Most SMTs did not appear to be stressed or under major pressure; they had high levels of confidence in their ability to succeed, and little pressure from stakeholders.

While this is an indication that SMTs are well-prepared for responding to oil spills, it is unclear whether this indication is reliable. In a real spill, teams would likely be operating under more pressure from the operator, from the public, and from government regulators. There would be no breaks, and participants would be tired and stressed. Research suggests that teams often operate differently under these conditions.

Our recommendation is that BSEE continue to conduct unannounced exercises and consider ways to introduce additional stress and pressure for some of these exercises. Adding more difficulties like equipment deployment and coordinating with stakeholders would help – BSEE already includes these in exercises but they could be included more often or be made more challenging. Other creative ways of adding stress could include varying the schedule (e.g., so spills occur in the middle of the night) or requiring exercises to run longer (e.g., 24 continuous hours or even a full week). Adding pressure would require creative crafting of pressure-inducing phenomena within the exercise. One example is to create an increase in perceived pressure from the public. This could be accomplished through a required press conference to which real media representatives (or actors) are invited, a single media representative with a video camera interviewing the IC (for the entire team to review later), a fake twitter feed, or even involving real members of the public to react as if the spill were real.

Perceived pressure from the government could also be created by the exercise designers, by including visits from high-level stakeholders (actors or real), adding injects requesting information or expressing frustration from government stakeholders, or ensuring that government representatives who are "playing" in the exercise express frustration or throw realistic roadblocks to the team. Finally, stress and pressure can be induced through realistic frustrations experienced by the team, such as an inability to get correct information: exercise designers could provide conflicting information as the exercise continues. These ideas can be used separately or together to increase stress and pressure.

5.2 Recommendations: Beyond the Current Program

The following recommendations may require changes to the current practices related to regulations and policy, so they may be more difficult to implement.

5.2.1   Investigate learning by contractors and operators

In the Gulf, most exercises are carried out by contracted SMTs. The operator supervises and allocates resources, but the response is managed by a contractor. The same few contractors work for most of the operators in the region. As a result, these contractors carry out many exercises, and the operators have limited involvement in them.

This situation may or may not represent a problem. Since operators are involved only in a supervisory role (except in source control exercises), they may not learn very much; if they were more involved, they might learn better ways to respond and update their response plans, for example. On the other hand, the SMTs get a large amount of practice executing various response plans for various operators, which might lead to greater expertise. An additional issue is that the parties have different incentives in an exercise: the operator wants to pass, and the SMT wants to impress the operator.

Our recommendation is that BSEE further investigate this situation to determine whether or not it poses a problem. Exercises may need to be designed differently to account for this situation.

5.2.2   Conduct some larger-scale exercises

BSEE's current SMT exercise program consists largely of small-scale table-top exercises (with separate equipment deployment drills). We have already recommended (above) that these table-top exercises be enhanced with stakeholder interactions, equipment deployment, and additional pressure and stress.

However, there are some elements that cannot be tested or are much more difficult to test in a small-scale exercise. These include (1) the management of a large response organization, (2) later parts of a spill response beyond the first few hours, (3) ongoing interactions among many stakeholders, (4) stress and tiredness due to long work-hours, and (5) unexpected events (because there is little time for them to arise in the current exercises).

Our recommendation is that BSEE evaluate some large-scale exercises, in which larger response teams are required, the response goes on for more than a few hours, and later parts of the response timeframe are played out. Doing so would enable the evaluation of the operator's ability to manage the five elements described above.

Larger-scale exercises are more expensive, so if resources are not available to conduct them, other opportunities may exist. Large-scale area exercises are already conducted,

and BSEE is already involved in these. Evaluating area exercises specifically with regard to the five elements described above is one resource-inexpensive way to accomplish this recommendation. A second idea is to incorporate some elements into BSEE's existing exercises, expanding them slightly in important dimensions. For example, some could simulate the response beginning on day 2 rather than day 1, while others could include some unexpected events.

5.2.3 Evaluate operators' ability to deal with situations outside the plan

Currently, BSEE is tasked with evaluating an operator's ability to execute its response plan. However, real oil spills have often presented issues or challenges that were not included in the plan. Therefore, the ability to execute a response plan is not enough to ensure an effective spill response.

Our recommendation is that BSEE consider evaluating an operator's ability to deal with situations outside the plan. This can be accomplished one of two ways. The first is to design exercise scenarios that contain challenges not included in the plan. This may be difficult, however, because we may not be able imagine such challenges. The second way to accomplish this recommendation is to evaluate an operator's ability to deal with unexpected situations in general. Our framework details several capabilities which, if mastered, are likely to help teams deal with new situations when they arise. Using the framework to evaluate these capabilities is another way to accomplish this recommendation.

This page intentionally left blank.

6 Ideas for Future Work

The following sections provide descriptions of avenues for future work that were suggested by our work on this project. Some are directly related to this project, while others are inspired by the work we have done without being directly related.

6.1 Future Work: Building on the Framework

The following ideas for future work directly build on the framework developed in this project.

6.1.1 Add levels of performance and difficulty to the framework

In its current state, the framework describes key capabilities, but it does not define levels of performance for each capability. For example, what constitutes excellent coordination, satisfactory coordination, or poor coordination? Future work could define levels of performance for each capability in the framework, based on the literature on each capability, observation of exercises, and discussions with subject matter experts.

Similarly, the current framework describes exercise design components that can be included in an exercise, but it does not define levels of difficulty for each component. For example, how would one define an exercise that makes coordination easy, moderately difficult, or very difficult? Future work could define levels of difficulty for each exercise design component, based on observation of exercises and discussions with subject matter experts and exercise participants.

6.1.2 Evaluate framework in action to refine measures and components

The framework was designed based on observation of exercises, but it has not been tested by the framework's intended users. Two elements of the framework would particularly benefit from such testing: the evaluation measures and the exercise design components.

The evaluation measures are intended to enable an exercise observer to easily evaluate the performance of a capability. A future research team could work with BSEE exercise evaluators to test how easily the measures can be observed and recorded in an exercise, and refine the measures accordingly.

The exercise design components are intended to enable an exercise designer to easily create and enhance an exercise scenario. A future research team could work with BSEE exercise designers to test how easily the components can be utilized, and refine the collection of components accordingly.

6.1.3 Determine which capabilities drive a successful response

The current framework outlines a large collection of capabilities that enable spill response, based on literature and observation of exercises. However, it is likely that some subset of these capabilities are primary drivers of successful response, while others contribute more indirectly. Future research could investigate which of the capabilities are the most important in enabling successful spill response. A first step would be to define, based on the literature, the features of a successful spill response; a second step would be to link capabilities to these features, based on observation of exercises, reports from past spills, discussion with subject matter experts, and, if possible, quantitative measurement.

6.2 Future Work: Exercises, Plans, and Preparedness

The following ideas for future work are inspired by our team's efforts in this project.

6.2.1 Investigate what operators learn from BSEE exercises

An exercise program is an important opportunity for improving response readiness through learning. Exercises can be used to evaluate whether an operator is prepared to respond to a spill, and they also uncover areas that require improvement.

However, it is not clear what exactly is learned by operators and other stakeholders, and whether those lessons are retained and acted upon. Without learning from the exercises, and without follow-up actions to implement those lessons, response readiness may not be improving.

Future researchers could investigate what is learned from exercises by each party (operators, SMT contractors, and other stakeholders). For example, the following questions could be investigated:
- Do participants/organizations improve their fundamental skills, such as operating software, writing appropriate documents, and hosting meetings?
- Do participants/organizations learn any new facts, such as identifying changes in processes for interacting with stakeholders, finding problems with the written response plan, or developing new ideas for source control methods?
- Do participants/organizations identify any broad areas for improvement, such as realizing that established coordination methods are inadequate, or that source control methods will not work for some types of spills?
- Are any of these "learnings" followed up by the participants/organizations in the months following the exercise? Are changes made in the organization's practices? Are new internal trainings implemented? Are response plans revised?

It would be best to investigate each of these questions at several times: (1) directly after the exercise, (2) three months after the exercise, and (3) one year after the exercise. In this manner, it would be possible to learn whether lessons persist and receive appropriate follow-up.

Investigating learning in this manner would also shed light on an important issue. In the Gulf region, one of the busiest regions, most of the exercise activities are carried out by SMT contractors rather than by the operators themselves. Are the contractors learning from each exercise, or are they so saturated by exercises that each one is not very useful? Are the operators learning anything from exercises, given their limited involvement in exercise activities? Such questions could be answered at least partially by this research idea.

6.2.2　Work with operators to evaluate and improve exercises

As described in the section above, exercises represent an important opportunity for operators to learn and to improve their readiness for response. While BSEE's teams receive informal feedback from operators during hotwash sessions, there is no formal process for asking the operators what could be done to make the exercises more useful to them. Such feedback could be useful for improving the exercise program.

A future research team could develop a formal process for requesting feedback from operators on the exercise program, to identify (1) useful aspects of exercises and (2) recommended changes. In one of the exercises our research team observed, the operator noted that the detail included in the source control portion of the exercises pushed their team to develop some useful new ideas, so the exercise was very helpful to them. It might be helpful to BSEE to know what design aspects of the exercises make them useful to operators. A first step would be to conduct interviews with key operators and SMT contractors, then to develop a more formal survey that could be used to solicit structured feedback from a variety of operators and other exercise participants.

6.2.3　Framework for evaluating response plans in many scenarios

It was often emphasized during this research project that the main goal of BSEE's exercise program was to determine whether an operator could execute its response plan. Evaluating the adequacy of the response plan itself is a separate but very important issue.

Response plans are difficult to evaluate because the oil spill event is uncertain. In many historical cases, response plans were inadequate because some aspects of the spill scenario had not been imagined, and therefore had not been planned for.

A future research team could build a "library" of spill scenarios, based on historical spill scenarios and based on ideas from a panel of experts about what future spills might look like. Such a library would enable BSEE to evaluate an operator's response plans for adequacy of response to a variety of different and challenging scenarios. Building such a library would require identifying and collecting key features of historical oil spills, then assembling an expert panel to imagine features of future oil spills based on new technologies and environmental conditions.

An additional important question is whether written plans are good enough or whether flexibility and improvisation are critical components of planning. Currently, an organization's ability to make changes to a plan in response to new events is not a critical element of exercise evaluation; instead, exercises are intended to demonstrate that an

organization can execute its plan as written. However, the ability to improvise around the plan when required might be a critical aspect of a response organization's skillset because it is difficult to plan for every possible scenario, especially in a rapidly changing technological setting.

This research project would enable BSEE to determine whether and how to evaluate improvisation as a part of response planning. For example, identifying which processes and skills enable organizations to improvise could enable BSEE to look for these processes and skills in response plans and in exercises.

7 References

Arnold, J. J., Johnson, L. M., Tucker, S. J., Malec, J. F., Henrickson, S. E., & Dunn, W. F. (2009). Evaluation Tools in Simulation Learning: Performance and Self-Efficacy in Emergency Response. *Clinical Simulation in Nursing*, *5*(1), e35–e43. http://doi.org/10.1016/j.ecns.2008.10.003

Beamon, B. M., & Balcik, B. (2008). Performance measurement in humanitarian relief chains. *International Journal of Public Sector Management*, *21*(1), 4–25. http://doi.org/10.1108/09513550810846087

Berlin, J. M., & Carlström, E. D. (2015). The Three-Level Collaboration Exercise - Impact of Learning and Usefulness. *Journal of Contingencies and Crisis Management*. http://doi.org/10.1111/1468-5973.12070

Biddinger, P. D., Cadigan, R. O., Auerbach, B. S., Burstein, J. L., Savoia, E., Stoto, M. A., & Koh, H. K. (2008). On linkages: using exercises to identify systems-level preparedness challenges. *Public Health Reports*, *123*(1), 96–101.

Brady, T. F. (2003). Public health: emergency management: capability analysis of critical incident response (pp. 1863–1867). Winter Simulation Conference.

Carroll, J. S., Rudolph, J. W., & Hatakenaka, S. (2002). Learning from experience in high-hazard organizations. *Research in Organizational Behavior*, *24*, 87–137. http://doi.org/10.1016/S0191-3085(02)24004-6

Cashman, D. (2011). Designing Operations-based Exercises based on Artificialities and Realism. In *2011 International Oil Spill Conference*. American Petroleum Institute.

Cashman, D., Stephens, J., & Boyles, L. T. (2003). Designing, Executing, and Evaluating a US Government-Led National Preparedness for Response Exercise Program (PREP) Drill. In *2003 International Oil Spill Conference* (pp. 597–602). American Petroleum Institute.

Chi, C.-H., Chao, W.-H., Chuang, C.-C., Tsai, M.-C., & Tsai, L.-M. (2001). Emergency medical technicians' disaster training by tabletop exercise. *The American Journal of Emergency Medicine*, *19*(5), 433–436. http://doi.org/10.1053/ajem.2001.24467

Cooper, S., Cant, R., Porter, J., Sellick, K., Somers, G., Kinsman, L., & Nestel, D. (2010). Rating medical emergency teamwork performance: Development of the Team Emergency Assessment Measure (TEAM). *Resuscitation*, *81*(4), 446–452. http://doi.org/10.1016/j.resuscitation.2009.11.027

Dausey, D. J., Buehler, J. W., & Lurie, N. (2007). Designing and conducting tabletop exercises to assess public health preparedness for manmade and naturally occurring biological threats. *BMC Public Health*, *7:92*. http://doi.org/10.1186/1471-2458-7-92

Descatha, A., Loeb, T., Dolveck, F., Goddet, N.-S., Poirier, V., & Baer, M. (2009). Use of Tabletop Exercise in Industrial Training Disaster. *Journal of Occupational and Environmental Medicine*, *51*(9), 990–991. http://doi.org/10.1097/JOM.0b013e3181b3265f

Fautua, D., Schatz, S., Reitz, E., & Bockelman, P. (2014). Institutionalizing Blended Learning into Joint Training: A Case Study and Ten Recommendations. In *Interservice/Industry Training, Simulation, and Education Conference 2014*.

Franks, N., Knutson, S., Parker, H. A., & LeJeune, F. (2011). Plan-holder exercises –

How to recapture a lost opportunity for area plan improvement. In *2011 International Oil Spill Conference*. American Petroleum Institute.

Gagné, R. M., Wager, W. W., Golas, K. C., & Keller, J. M. (2005). *Principles of instructional design* (5th ed.). Wiley Online Library. http://doi.org/10.1002/pfi.4140440211

Gebbie, K. M., Valas, J., Merrill, J., & Morse, S. (2006). Role of exercises and drills in the evaluation of public health in emergency response. *Prehospital and Disaster Medicine*, *21*(03), 173–182.

Harrald, J. (2006, March). *Testimony for the Senate Homeland Security Government Affairs Committee on Hurricane Katrina: Recommendations for Reform*.

Jackson, B. A., & McKay, S. (2011). Preparedness Exercises 2.0 Alternative Approaches to Exercise Design That Could Make Them More Useful for Evaluating and Strengthening Preparedness. *Homeland Security Affairs*, *7*.

Langley, A. (1999). Strategies for theorizing from process data. *Academy of Management Review*, *24*(4), 691–710. http://doi.org/10.5465/AMR.1999.2553248

Miles, M. B., & Huberman, a M. (1984). Qualitative Data Analysis: A Sourcebook of New Methods. *Sage Publications*.

National Preparedness for Response Exercise Guidelines (DRAFT). (2015, March). USCG, EPA & BSEE.

National Preparedness for Response Exercise Program Guidelines. (2002, August). USCG, EPA, & DOI.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

Rosen, M., Weaver, S., Lazzara, E., Salas, E., Wu, T., Silvestri, S., … King, H. (2010). Tools for evaluating team performance in simulation-based training. *Journal of Emergencies, Trauma, and Shock*, *3*(4), 353. http://doi.org/10.4103/0974-2700.70746

Silenas, R., Akins, R., Parrish, A. R., & Edwards, J. C. (2008). Developing disaster preparedness competence: An experiential learning exercise for multiprofessional education. *Teaching and Learning in Medicine: An International Journal*, *20*(1), 62–68. http://doi.org/10.1080/10401330701798311

Target Capabilities List: A companion to the National Preparedness Guidelines. (2007, September). U.S. Department of Homeland Security.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

U.S. Department of Homeland Security. (2013). *Homeland Security Exercise and Evaluation Program*.

Weick, K. E. (1993). The Collapse of Sensemaking in Organizations: The Mann Gulch Disaster. *Administrative Science Quarterly*, *38*(4), 628. http://doi.org/10.2307/2393339

Zhou, B., Sun, G., Zhang, X., Xu, J., Lai, J., Du, X., … Kimura, R. (2015). Development of Web-Based Tabletop Emergency Earthquake Exercise System. *Journal of Disaster Research*, *10*(2), 217–224.

8 APPENDIX A: Capabilities

8.1 Functional Capabilities

8.1.1 Deploy and Demonstrate Use of Equipment

The ability to successfully move technical response resources to the affected area and demonstrate the use of resources for their intended purpose.

Evaluation Measures and Techniques

Measure 1 (Output)	
Responders *deployed* the appropriate response equipment.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders appropriately and safely *demonstrated* the use of response equipment. If they were unable to perform, they discussed how they would operate the equipment.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Deployment and Demonstration of Equipment	Require responders to demonstrate ability to use equipment
Require Strategic, Operational or Tactical Decisions	Challenges responders to consider best appropriate action, such as which equipment to deploy and use
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Prompts responders to coordinate efforts with other on-scene responders
Incorporate Time Pressure	Challenges responders to execute capability under pressure

Relevant References

Drills & Exercises Evaluation Guidance Manual. (2010, February). California Department of Fish and Game, Office of Spill Prevention and Response.

National Preparedness for Response Exercise Guidelines (DRAFT). (2015, March). USCG, EPA & BSEE.

National Preparedness for Response Exercise Program Guidelines. (2002, August). USCG, EPA, & DOI.

8.1.2　Ensure Security and Safety

Description

Ensure the security and safety of the general public and provide for a secure and safe work environment for oil spill responders. This includes establishing a strong safety culture and creating the appropriate processes and protocols to ensure protection from immediate (e.g., accidents) and long-term hazards (e.g., exposure to oil fumes). This may include activities such as securing the incident and cleanup sites, establishing no-fly zones, conducting or coordinating air monitoring, requiring pre-work safety briefings, or developing safety plans and "stop work" protocols.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders considered how to ensure *responders* were operating in a safe and secure environment when making decisions, taking action or developing plans.	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 2 (Process)	
Responders considered how to ensure *the public* was safe from the effects of the spill when making decisions, taking action or developing plans.	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 3 (Output)	
Responders created a safety and security plan.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 4 (Outcome)	
Responders were able to ensure *responders* operated in a safe and secure environment.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 5 (Outcome)	
Responders were able to ensure *the public* was safe from the effects of the spill.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Complex coordination challenges ability to maintain security and safety. (e.g., who is approved to be in restricted area?)
Require Strategic, Operational or Tactical Decisions	Should night operations be conducted?
Require Plans or Strategies	What is the Safety Plan?
Incorporate Unexpected Situations	Fishing vessel is in restricted area
Moderate Relevance	
Require Updates, Reports or Information	What is the status of all evacuated and injured personnel?
Incorporate Changes in the Situation	High heat is expected the next two days
Incorporate Unclear Situations or Problems	Report of a many injuries with no detail
Require Meetings	Require a safety planning meeting
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Many fishing vessels are contaminated with the spilled oil
Incorporate Multiple Events	Many responders on the water are passing out due to heat exhaustion and sea conditions are becoming rough
Incorporate Resource Insufficiencies	There are not enough MedEvac helicopters
Weak Relevance	
Incorporate Incomplete or Conflicting Information	Multiple people reporting an injury event as different injury events, causing confusion over ground truth
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Only have enough decontamination suites for X responders
Require Execution of Plans, Procedures or Processes	Execute perimeter security plan

Incorporate Public Perception of Priorities, Goals or Tasks	Public is watching oil come ashore and are starting to clean it up by themselves
Require Authorizations	Safety plan requires approval by Unified Command
Require Group Decision Making	Under what conditions can different responders operate if there is high heat, rough seas, limited personnel, etc.
Incorporate Ambiguous "Correct" Decisions	Which MedEvac company should we go with?
Incorporate Events or Tasks that Exceed Experience or Expertise of Responders	Environmental stress of rough seas or heat

Relevant References

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

Target Capabilities List: A companion to the National Preparedness Guidelines. (2007, September). U.S. Department of Homeland Security.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

8.1.2.1 Execute Emergency Procedures (Sub-Capability)

The ability to recognize the need for and execute pre-defined emergency procedures to ensure the safety of personnel and minimize impact of the situation (e.g., emergency shut-down).

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders *identified* the appropriate or relevant emergency procedures to be used in a particular situation.	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 2 (Output)	
Responders safely and correctly *executed* the emergency procedures.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation	**Comments:**

8.1.3   Inform and Warn the Public

Description
Create and deliver coordinated and actionable information to the public in a timely manner. Information should be appropriate to the target audience, including information on the dangers associated with the spill as well as updates on response and recovery efforts. Additionally, information should appropriately manage the expectations of the public and be delivered through channels easily accessible by the public.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders *created* actionable information that was appropriate for the public audience(s).	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output/Outcome)	
Responders *delivered* information that was accessible, useful, timely, clear, and sufficiently comprehensive to the public.	**Result:** 1. Unsatisfactory 2. Satisfactory 3. Very Satisfactory 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Complex coordination challenges ability to maintain security and safety. (e.g., who is approved to be in restricted area?)
Require Strategic, Operational or Tactical Decisions	Should night operations be conducted?
Require Plans or Strategies	What is the Safety Plan?
Incorporate Unexpected Situations	Fishing vessel is in restricted area
Moderate Relevance	
Require Updates, Reports or Information	What is the status of all evacuated and injured personnel?
Incorporate Changes in the Situation	High heat is expected the next two days
Incorporate Unclear Situations or Problems	Report of a many injuries with no detail
Require Meetings	Require a safety planning meeting
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Many fishing vessels are contaminated with the spilled oil
Incorporate Multiple Events	Many responders on the water are passing out due to heat

	exhaustion and sea conditions are becoming rough
Incorporate Resource Insufficiencies	There are not enough MedEvac helicopters
Weak Relevance	
Incorporate Incomplete or Conflicting Information	Multiple people reporting an injury event as different injury events, causing confusion over ground truth
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Only have enough decontamination suites for X responders
Require Execution of Plans, Procedures or Processes	Execute perimeter security plan
Incorporate Public Perception of Priorities, Goals or Tasks	Public is watching oil come ashore and are starting to clean it up by themselves
Require Authorizations	Safety plan requires approval by Unified Command
Require Group Decision Making	Under what conditions can different responders operate if there is high heat, rough seas, limited personnel, etc.
Incorporate Ambiguous "Correct" Decisions	Which MedEvac company should we go with?
Incorporate Events or Tasks that Exceed Experience or Expertise of Responders	Environmental stress of rough seas or heat

Relevant References

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

Target Capabilities List: A companion to the National Preparedness Guidelines. (2007, September). U.S. Department of Homeland Security.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

8.1.4   Notify and Update Stakeholders and Response Partners

Description
The ability to notify stakeholders and response partners of the oil spill and regularly update stakeholders and partners with essential information during the response. This capability differs from Inform and Warn the Public as it specifically relates to key stakeholders (e.g., elected officials in affected locations, regulators, U.S. Coast Guard, etc.) and response partners (e.g., wildlife response organizations, U.S. Coast Guard, dispersant providers, etc.).

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders notified all the stakeholders, response partners listed in their plan(s), and others involved in the response of the oil spill.	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 2 (Process)	
Responders regularly provided essential information and updates to all stakeholders and response partners involved in the response.	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 3 (Outcome)	
Stakeholders and response partners felt adequately informed and communicated with during the response.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Execution of Plans, Procedures or Processes	EX: Require execution of notification plan
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Stakeholder requests regular updates
Deployment and Demonstration of Equipment	EX: Use communications equipment to notify

Relevant References

National Preparedness for Response Exercise Guidelines (DRAFT). (2015, March). USCG, EPA & BSEE.

National Preparedness for Response Exercise Program Guidelines. (2002, August). USCG, EPA, & DOI.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

8.1.5　Contain Spill to Mitigate Environmental Impact

Description

Stop the physical spread and impact of the oil spill by taking intentional and proactive actions to contain the spill in a timely manner. For oil spills, a primary objective of this capability is to prevent the spill from reaching the shoreline and impacting the environment. A secondary objective may be to corral the spill so it is easier to clean up.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders created a plan or strategy that prioritized containment efforts and methods to mitigate environmental impact.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders deployed and demonstrated the use appropriate resources (e.g., boom, skimmers, personnel, etc.) that would contain the oil and mitigate environmental impact.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 3 (Outcome)	
In the judgement of regulator technical specialists or evaluators, the decisions made and actions taken would be successful in quickly containing and mitigating environmental impact.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Baseline Event Specifies Precipitating Event, Event Impact and Location, and Environmental Conditions	Prompts responders to begin making decisions and taking actions to respond to event.
Require Updates, Reports or Information	EX: Require responders to develop a report on status of spill containment and environmental impact
Require Strategic, Operational or Tactical Decisions	EX: Require decision on where to contain spill first
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Prompts responders to prioritize where to contain spill first given the resources available
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Prompts responders to coordinate spill containment efforts with other aspects of the response
Moderate Relevance	
Require Group Decision Making	Prompts responders to discuss spill problems and make decisions as a group
Incorporate Incomplete or Conflicting Information	Challenges responders to contain spill even though all locations may not be identified
Require Plans or Strategies	Ex: Require spill containment strategy
Require Authorizations	Challenges responders to quickly obtain needed authorizations for specific approaches or strategies for containing the spill
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Challenges responders to contain spill in such a way that it also reflects priorities and goals of stakeholders
Require Meetings	EX: Require a spill containment strategy meeting
Weak Relevance	
Incorporate Resource Insufficiencies	Challenges responders to contain spill without all needed resources available
Incorporate Changes in the Situation	Challenges responders to continue to contain spill despite changes to situation (e.g., weather, roving slick, etc.)
Incorporate Unexpected Situations	Challenges responders to continue to contain spill despite an unexpected situation (e.g., high heat)
Incorporate Ambiguous "Correct" Decisions	Challenges responders to develop a "good" plan or strategy for spill containment despite no clear best way
Incorporate Public Perception of Priorities, Goals or Tasks	Challenges responders to consider public's perception of spill containment priorities, goals or tasks
Incorporate Multiple Events	Challenges responders to contain spill when multiple events are occurring
Incorporate Events or Tasks that Stress Plans	Challenges responders to develop new plans or strategies for containing spill
Require Deployment and Demonstration of Equipment	Prompts responders to contain spill by deploying and demonstrating use of specific equipment
Require Execution of Plans, Procedures or Processes	Prompts responders to execute a plan, procedure or process for containing the spill
Incorporate Ambiguous Group Structures and Operating Norms	Challenges responders to contain the spill despite an ambiguous organizational structure or operating norms
Incorporate Plan or Strategy Conflicts	Challenges responders to de-conflict plans or strategies (e.g., ACP conflicts with GRPs for containment strategies)

Relevant References

Fingas, M. F., Duval, W. S., & Stevenson, G. B. (1979). *The basics of oil spill cleanup: with particular reference to southern Canada*. Quebec: Environmental Emergency Branch, Environmental Protection Service, Environment Canada.

The Exxon Valdez Oil Spills - Final Report, State of Alaska Response. (1993, June). Alaska Department of Environmental Conservation.

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

Pearson, C. M., & Mitroff, I. I. (1993). From crisis prone to crisis prepared-a framework for crisis management.pdf. *Academy of Management Executive*, *7*(1), 48–59.

Target Capabilities List: A companion to the National Preparedness Guidelines. (2007, September). U.S. Department of Homeland Security.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

8.1.6   Mitigate Economic Impact of Spill

Description
Responders should take intentional and proactive actions to mitigate the economic impact of the oil spill. Oil spills can greatly impact livelihoods and industries and responders should begin to address these concerns during a response. For example, responders may establish and execute activities such as the Vessel of Opportunity program to reduce adverse impact on the affected public's income and livelihood.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders identified affected stakeholders and their long-term interests or needs in recovering from the spill.	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 2 (Process)	
Responders developed plans or established programs to mitigate the economic impact of the spill (e.g., vessel of opportunity program, employing local residents for cleanup, establishing claims fund and procedures, etc.).	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 3 (Output)	
Stakeholders and the public felt satisfied that adequate measures were being taken to mitigate the economic impact of the spill.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Ambiguous "Correct" Decisions	Unknown if decisions will mitigate impact as intended
Require Strategic, Operational or Tactical Decisions	EX: Require decision to establish a Vessel of Opportunity program
Incorporate Incomplete or Conflicting Information	It is difficult to predict/project economic impacts of interventions
Require Plans or Strategies	EX: Require an economic mitigation plan
Incorporate Plan or Strategy Conflicts	Economic mitigation strategy may conflict with response priorities and strategies
Incorporate Events or Tasks that Stress Plans	Challenges responders to develop additional plans

8.1.7   Control and Stop Oil Spill at Source

Description

Stop the active spilling of oil by taking intentional and proactive actions that stop or reduce the amount of spillage in a timely manner.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders mobilized the appropriate resources early in the response to help control the flow of the spill at the source before a permanent solution was implemented. This may have included reducing the flow rate or diverting flow to a containment vessel or area.	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 2 (Process)	
Responders developed a strategy and mobilized the appropriate resources to permanently stop the spill at the source in a timely manner.	**Result:** Yes or No
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 3 (Output)	
Regulator technical specialists felt that the decisions made and actions taken would be successful in quickly controlling the spill and permanently stopping the spill at the source.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Prompts responders to prioritize source control efforts
Incorporate Resource Insufficiencies	Challenges responders to prioritize source control efforts without all resources available
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Prompts responders to coordinate source control efforts with other aspects of the response
Require Plans or Strategies	EX: Require a flow control strategy
Require Deployment and Demonstration of Equipment	Prompts responders to demonstrate proficiency with specific source control activities
Require Execution of Plans, Procedures or Processes	EX: Require execution of flow control strategy

Relevant References

National Preparedness for Response Exercise Guidelines (DRAFT). (2015, March). USCG, EPA & BSEE.

National Preparedness for Response Exercise Program Guidelines. (2002, August). USCG, EPA, & DOI.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

8.1.8   Fight Fires in a Marine Environment

Description
The ability to assess the need for and deliver firefighting capabilities that manage and extinguish marine-based fires resulting from the oil spill or subsequent response activities in a timely manner.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders, early in the response, assessed and ordered the quantity and types of resources needed for marine firefighting capabilities.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders prevented and/or reduced the risk of unintentional marine-based fires.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Deployment and Demonstration of Equipment	EX: Require demonstration of water pumping equipment
Require Execution of Plans, Procedures or Processes	EX: Require execution of firefighting plan
Require Strategic, Operational or Tactical Decisions	Challenges responders to consider best appropriate action
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Prompts responders to coordinate firefighting efforts with other on-scene responders

Relevant References
Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

National Preparedness for Response Exercise Guidelines (DRAFT). (2015, March). USCG, EPA & BSEE.

National Preparedness for Response Exercise Program Guidelines. (2002, August). USCG, EPA, & DOI.

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

Target Capabilities List: A companion to the National Preparedness Guidelines. (2007, September). U.S. Department of Homeland Security.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

8.1.9　Conduct Human Search and Rescue

Description

Deliver search and rescue capabilities to find and rescue people affected by the oil spill with the goal to save the greatest number of endangered lives in the shortest time possible.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders developed search and rescue plans or strategies to rescue people affected by the oil spill.	**Result:** Yes or No

Data collection methods:
   • Observation
   • Documentation
	Comments:

Measure 2 (Output)	
Responders' plans or strategies included anticipated search grids, likely locations of people in need of rescue, and safety/rescue protocols (e.g., handling injured people, coordination, or transportation points).	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well

Data collection methods:
   • Observation
   • Documentation
   • Participant Hotwash
   • Controller/Evaluator Debriefing
	Comments:

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Ambiguous "Correct" Decisions	EX: Require responders to decide on search pattern
Require Strategic, Operational or Tactical Decisions	EX: Require decision on what crews to send out, when and where
Incorporate Incomplete or Conflicting Information	Challenges responders to develop search patterns based on where people/wildlife might be
Require Plans or Strategies	EX: Require a search and rescue plan
Incorporate Events or Tasks that Stress Plans	Challenges responders to develop additional plans

Relevant References

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

Target Capabilities List: A companion to the National Preparedness Guidelines. (2007, September). U.S. Department of Homeland Security.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

8.1.10   Conduct Wildlife Search and Rescue

Description
Deliver search and rescue capabilities to find, rescue and decontaminate wildlife affected by the oil spill with the goal to save the greatest number of endangered lives in the shortest time possible.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders developed search and rescue plans or strategies to rescue wildlife affected by the oil spill.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders' plans or strategies included anticipated search grids, likely locations of wildlife in need of rescue, and safety/rescue protocols (e.g., handling oiled wildlife, coordination, or transportation points).	**Result:** 4. Did Not Perform Well 5. Performed Satisfactorily 6. Performed Well
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Ambiguous "Correct" Decisions	EX: Require responders to decide on search pattern
Require Strategic, Operational or Tactical Decisions	EX: Require decision on what crews to send out, when and where
Incorporate Incomplete or Conflicting Information	Challenges responders to develop search patterns based on where people/wildlife might be
Require Plans or Strategies	EX: Require a search and rescue plan
Incorporate Events or Tasks that Stress Plans	Challenges responders to develop additional plans

Relevant References

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

Target Capabilities List: A companion to the National Preparedness Guidelines. (2007, September). U.S. Department of Homeland Security.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

8.1.11 Conduct Shoreline Cleanup and Restoration

Description

The cleanup and initiation of restoration planning for the land environment impacted by the marine oil spill. This includes the removal of oil and decontamination of the affected wildlife and land area. This may also include recovery actions to return the affected land to its pre-spill state or a better state.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders initiated planning for long-term restoration activities that would return the affected shoreline to its pre-spill or better condition.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 2 (Process)	
Responders created the appropriate plans and deployed the appropriate resources (e.g., OSROs, wildlife partner organizations, protective gear, etc.) to clean up the oil, and capture, decontaminate, and treat the affected wildlife. Responders recognized each of these needs and took appropriate steps to begin addressing them in coordination with other plans and activities.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 3 (Output)	
Responders were able to clean up the affected shoreline and capture, decontaminate, and treat the affected wildlife.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Strategic, Operational or Tactical Decisions	<u>EX:</u> Require decision on which shoreline to cleanup first
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Challenges responders to manage others' cleanup priorities with the resources available
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Prompts responders to coordinate shoreline cleanup efforts with other aspects of the response
Moderate Relevance	
Incorporate Incomplete or Conflicting Information	Challenges responders to cleanup shoreline even though all locations may not be identified
Incorporate Changes in the Situation	Challenges responders to continue to cleanup shoreline despite changes to situation (e.g., weather, roving slick, etc.)
Incorporate Public Perception of Priorities, Goals or Tasks	Challenges responders to consider public perception or priorities in cleanup efforts
Require Plans or Strategies	<u>EX:</u> Require shoreline cleanup strategy
Incorporate Events or Tasks that Stress Plans	Challenges responders develop new plans or strategies for shorelines cleanup
Incorporate Ambiguous "Correct" Decisions	Challenges responders to develop a "good" plan or strategy for shoreline cleanup despite no clear best way
Require Authorizations	Challenges responders to use to plan and use acceptable cleanup methods
Weak Relevance	
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Prompts responders to prioritize what gets cleaned up first given the resources available
Incorporate Resource Insufficiencies	Challenges responders to manage cleanup priorities without all needed resources available
Require Meetings	<u>EX:</u> Require shoreline cleanup strategy meeting
Incorporate Unexpected Situations	Challenges responders to continue to cleanup shoreline despite an unexpected situation (e.g., workforce sick with flu)
Incorporate Events or Tasks that Vary in Size and Require Different Response Levels	Challenges responders to cleanup any size cleanup

Relevant References

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

The Exxon Valdez Oil Spills - Final Report, State of Alaska Response. (1993, June). Alaska Department of Environmental Conservation.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

8.1.12   Conduct Marine Cleanup and Restoration

Description

The cleanup and initiation of restoration planning for the marine environment impacted by the oil spill. This includes the removal of oil and decontamination of affected the wildlife and ocean area. This may also include long-term recovery actions to return the affected ocean area to its pre-spill state or a better state.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders initiated planning for long-term restoration activities that would return the affected marine environments and habits to their pre-spill or better condition.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 2 (Process)	
Responders created the appropriate plans and deployed the appropriate resources (e.g., OSROs, wildlife partner organizations, protective gear, etc.) to clean up the oil, capture, decontaminate, and treat the affected wildlife, decontaminate the affected vessels, and restore the affected habitats to their pre-spill or better conditions. Responders recognized each of these needs and took appropriate steps to begin addressing them in coordination with other plans and activities.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 3 (Output)	
Responders were able to clean up the oil on the water and capture, decontaminate, and treat the affected wildlife as well as decontaminate affected vessels.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Incomplete or Conflicting Information	Conflicting information is received on which area should be cleaned first
Incorporate Public Perception of Priorities, Goals or Tasks	Public has taken pictures of the response vessels still at the dock and are angry they are not yet deployed for cleanup
Incorporate Changes in the Situation	The seas are getting rougher
Incorporate Unexpected Situations	Skimming vessel loses power
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	There are several slick areas that will be cleaned up simultaneously be different crews
Require Deployment and Demonstration of Equipment	Practice vessel-based and aerial dispersant deployment
Moderate Relevance	
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Fisherman are frustrated they can't help cleanup
Incorporate Events or Tasks that Stress Plans	More oil than anticipated in plan has leaded
Incorporate Ambiguous Group Structures and Operating Norms	Oil slick has spread more widely than anticipated (e.g., possibly requiring several new cleanup Task Forces)
Incorporate Plan or Strategy Conflicts	Geographic Response Plan calls for skimming in a particular area, but dispersants are being used in surrounding area
Weak Relevance	
Require Group Decision Making	Where should we deploy cleanup boat crews?
Require Execution of Plans, Procedures or Processes	Execute cleanup strategy
Incorporate Resource Insufficiencies	Geographic Response Plan calls for booming, but there is no more boom available
Incorporate Ambiguous "Correct" Decisions	Do we skim, boom, deploy dispersants or conduct in-situ burning?
Incorporate Unclear Situations or Problems	Skimming equipment is not picking up oil as intended (e.g., must investigate/troubleshoot)

Relevant References

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

The Exxon Valdez Oil Spills - Final Report, State of Alaska Response. (1993, June). Alaska Department of Environmental Conservation.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

8.1.13   Conduct Salvage and Lightering

Description

The rescue of a wrecked ship and its cargo (salvage) from sea as well as the transfer of cargo and oil (lightering) from the wrecked ship to another ship or holding container.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders deployed the appropriate resources to 1) stabilize the vessel or rig to prevent additional spillage and maintain safety, 2) remove cargo, and 3) remove vessel or rig from spill area so it can be salvaged.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders safely 1) stabilized the vessel or rig to prevent additional spillage and maintain safety, 2) removed cargo, and 3) removed vessel or rig from spill area so it could be salvaged.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Deployment and Demonstration of Equipment	EX: Require demonstration of lightering equipment (e.g., cranes, pumps, etc.)
Require Execution of Plans, Procedures or Processes	EX: Require execution of salvage and lightering plan
Require Strategic, Operational or Tactical Decisions	Challenges responders to consider best appropriate action
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Prompts responders to coordinate salvage and lightering efforts with other on-scene responders

Relevant References

U.S. Coast Guard Incident Management Handbook. (2001). U.S. Coast Guard.

8.1.14   Separate and Dispose of Oil and Debris

Description
The ability to separate the oil/water/debris mixture and properly dispose of the collected material.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders obtained appropriate permits to separate and dispose of oil and debris.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders appropriately and safely ***separated*** the oil, water, and debris mixture. If they were unable to perform, they developed a suitable plan and deployed appropriate resources.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 3 (Output)	
Responders appropriately and safely ***disposed*** of the oil, water, and debris. If they were unable to perform, they developed a suitable plan and deployed appropriate resources.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Deployment and Demonstration of Equipment	EX: Require demonstration of oil and debris separation/disposal
Require Execution of Plans, Procedures or Processes	EX: Require execution of oil and debris management plan
Require Strategic, Operational or Tactical Decisions	Challenges responders to consider best appropriate action
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Prompts responders to coordinate separation and disposal efforts with other on-scene responders
Incorporate Time Pressure	Challenges responders to execute capability under pressure

Relevant References

Fingas, M. F., Duval, W. S., & Stevenson, G. B. (1979). *The basics of oil spill cleanup: with particular reference to southern Canada*. Quebec: Environmental Emergency Branch, Environmental Protection Service, Environment Canada.

U.S. Coast Guard Incident Management Handbook. (2001). U.S. Coast Guard.

8.2 Management and Support Capabilities

8.2.1　Set and Monitor Progress toward Goals and Objectives

Description of Capability
The ability to set strategic, operational, and tactical goals and objectives. The team should be able to identify and prioritize goals and objectives as well as align them with other response teams' goals and objectives. Goals and objectives should be specific and measureable. The team should also be able to monitor and track progress toward goals and objectives, including lower-level tasks and action items. This serves as a feedback mechanism to understand response effectiveness and to consider whether or not the response is functioning as intended.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders set meaningful and realistic goals and objectives given the situation/scenario.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Process)	
Responders established a reporting system to collect data relevant to measures.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 3 (Outcome)	
Responders set and tracked progress toward goals and objectives in such a way that it clearly communicated the priorities and effectiveness of the response.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Meetings	EX: Require meeting that reviews strategy and objectives
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Prompts need to set short-term and strategic goals and objectives to help manage response
Require Updates, Reports or Information	EX: Require a status report of progress toward goals and objectives
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Prompts need to set short-term and strategic goals and objectives that communicate priorities
Require Plans or Strategies	EX: Require an Incident Action Plan with short-term objectives
Incorporate Events or Tasks that Vary in Size and Require Different Response Levels	Size or complexity of scenario prompts need to set goals/objectives
Incorporate Incomplete or Conflicting Information	Challenges ability to set meaningful goals and objectives
Require Strategic, Operational or Tactical Decisions	EX: By when should the new Incident Command Post be established and everyone transferred to the new location?
Incorporate Resource Insufficiencies	Challenges ability to meet goals and objectives
Moderate Relevance	
Incorporate Ambiguous "Correct" Decisions	EX: Set appropriate goals and objectives for the complex and dynamic situation
Incorporate Unclear Situations or Problems	EX: Unknown cause or impact of spill yet, but goals and objectives need to be set
Weak Relevance	
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Challenges ability to meet goals and objectives
Incorporate Ambiguous Group Structures and Operating Norms	Challenges ability to meet goals and objectives
Incorporate Changes in the Situation	Challenges ability to meet goals and objectives
Incorporate Public Perception of Priorities, Goals or Tasks	Challenges ability to meet goals and objectives
Require Authorizations	Challenges ability to meet goals and objectives

Relevant References

DeShon, R. P., Kozlowski, S. W. J., Schmidt, A. M., Milner, K. R., & Wiechmann, D. (2004). A Multiple-Goal, Multilevel Model of Feedback Effects on the Regulation of Individual and Team Performance. *Journal of Applied Psychology*, *89*(6), 1035–1056. http://doi.org/10.1037/0021-9010.89.6.1035

Kline, T., & McGrath, J.-L. (1998). Development and validation of five criteria for evaluating team performance. *Organization Development Journal*, *16*(3), 19–27.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A Temporally Based Framework and Taxonomy of Team Processes. *The Academy of Management Review*, *26*(3), 356–376. http://doi.org/10.2307/259182

8.2.2　Conduct Strategy and Response Planning

Description

The ability to develop tactics that specify how different activities, responsibilities and resources will be used to achieve the goals and objectives. This capability includes the ability to plan the approach as well as to plan for contingencies if the approach is not successful or if the evolving situation necessitates a different approach. There is a complementary relationship between this capability and the capabilities listed under the Functional Capabilities category because they are dependent on effective response strategies and plans. Additionally, this capability differs from Coordinate Operations in that it focuses on planning the execution rather than executing.

Evaluation Measures and Techniques

Measure 1 (Output)	
Responders developed informal and formal plans or strategies during the response that were clear, timely and comprehensive.	**Result:** Yes or No
Data collection methods: 　• Observation 　• Documentation	**Comments:**

Measure 2 (Outcome)	
Responders developed plans and strategies during the response that met the needs of the response and did not conflict with each other.	**Result:** 1.　Did Not Perform Well 2.　Performed Satisfactorily 3.　Performed Well
Data collection methods: 　• Observation 　• Documentation 　• Participant Hotwash 　• Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Plans or Strategies	EX: Require responders to develop plan or strategy to address resources at risk.
Require Strategic, Operational or Tactical Decisions	Prompts responders to develop plans or strategies to support decisions
Require Updates, Reports or Information	EX: Require responders to develop a report that provides an update on the status of plans in development or completed
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Challenges responders to prioritize plan or strategy development efforts
Require Group Decision Making	Prompts responders to discuss and make decisions as a group about plans or strategies for response efforts
Incorporate Incomplete or Conflicting Information	Challenges responders to develop plans or strategies with incomplete or conflicting information
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Challenges responders to coordinate plans and strategies
Require Authorizations	Prompts responders to seek and manage approval process for plans or strategies
Incorporate Changes in the Situation	Challenges responders build flexible or modify plans or strategies based on the situation
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Challenges responders to consider stakeholder's priorities and goals when creating plans or strategies
Moderate Relevance	
Incorporate Ambiguous Group Structures and Operating Norms	Challenges responders to create plans or strategies despite ambiguous group structures or operating norms
Require Meetings	Prompts responders to discuss plans or strategies
Incorporate Plan or Strategy Conflicts	Challenges responders to de-conflict plans or strategies (e.g., 2 GRPs with A priority areas)
Require Execution of Plans, Procedures or Processes	Prompts responders to create a plan, procedure or process if none exists
Incorporate Public Perception of Priorities, Goals or Tasks	Challenges responders to consider the public's perception of response priorities, goals or tasks within plans or strategies
Incorporate Multiple Events	Challenges responders to prioritize planning efforts
Incorporate Unexpected Situations	Challenges responders build flexible or modify plans or strategies based on the situation
Weak Relevance	
Incorporate Resource Insufficiencies	Challenges responders to plan or strategize despite needed resource(s) being unavailable
Incorporate Interpersonal Conflicts	Challenges effectiveness of group planning or strategizing
Incorporate Events or Tasks that Stress Plans	Challenges responders to plan or strategize about response in absence of adequate prior plans
Require Deployment and Demonstration of Equipment	Prompts responders to plan or strategize about how to deploy and demonstrate the use of specific equipment
Incorporate Time Pressure	Challenges responders to plan or strategize about response with the pressure of time

Relevant References

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

LePine, J. A., Piccolo, R. F., Jackson, C. L., Mathieu, J. E., & Saul, J. R. (2008). A meta-analysis of teamwork processes: tests of a multidimensional model and relationships with team effectiveness criteria. *Personnel Psychology*, *61*(2), 273–307.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A Temporally Based Framework and Taxonomy of Team Processes. *The Academy of Management Review*, *26*(3), 356–376. http://doi.org/10.2307/259182

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

Target Capabilities List: A companion to the National Preparedness Guidelines. (2007, September). U.S. Department of Homeland Security.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

8.2.3 Assess and Monitor the Evolving Situation

Description

The ability to create and maintain an accurate and comprehensive understanding of the situation, including damage, needs, and priorities. This includes assessing and understanding the current situation as well as anticipating future situation(s) based on the information available. For example, responders may assess and monitor threats/hazards to operations, ongoing response activities, or public/stakeholder perception of the response. The Manage and Share Information Capability supports this capability in that it helps provide critical information about an evolving situation.

Evaluation Measures and Techniques

Measure 1 (Output)	
Responders created and shared information products that were timely and reflected an accurate and comprehensive understanding of the current and future situation.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 2 (Outcome)	
Responders developed an accurate and comprehensive understanding of the current and future situation positively impacts response operations (e.g., better coordination, quicker decision making, more informed decisions, etc.).	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Updates, Reports or Information	Challenges responders to accurately report on situation
Incorporate Incomplete or Conflicting Information	Challenges responders to accurately understand situation
Require Strategic, Operational or Tactical Decisions	Prompts responders to make decision based on assessment of situation
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Prompts responders to coordinate with others to develop an accurate understanding of situation
Incorporate Changes in the Situation	Challenges responders to identify significant changes to a situation that require attention
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Challenges responders to prioritize efforts based on the assessment of the situation
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Challenges responders to maintain awareness of stakeholders' impact on the situation
Moderate Relevance	
Require Meetings	Prompts responders to share information about the situation
Require Plans or Strategies	Prompts responders to use situation assessment information for planning
Require Group Decision Making	Prompts responders to have a shared understanding of the situation to make decisions
Incorporate Resource Insufficiencies	Challenges responders to understanding situation needs vs. current capabilities
Require Authorizations	Prompts responders to have an accurate assessment of situation to approve plans and reports
Incorporate Ambiguous "Correct" Decisions	Challenges responders to create a shared understanding of the situation to make decisions
Incorporate Public Perception of Priorities, Goals or Tasks	Challenges responders to maintain awareness of public's impact on the situation
Incorporate Multiple Events	Challenges responders' ability to assess and monitor the situation
Incorporate Unexpected Situations	Challenges responders' ability to assess and monitor the situation
Weak Relevance	
Incorporate Time Pressure	Challenges responders to accurately assess situation in a limited timeframe
Incorporate Unclear Situations or Problems	Challenges responders to have an accurate understanding of the situation to address unclear problems
Incorporate Events or Tasks that Stress Plans	Challenges responders to think independently about how they should assess and monitor the situation
Require Execution of Plans, Procedures or Processes	EX: Require responders to follow damage assessment plan
Incorporate Events or Tasks that Exceed Experience or Expertise of Responders	Challenges responders to accurately assess situation despite the stress of the situation
Incorporate Interpersonal Conflicts	Challenges responders to overcome interpersonal conflict to accurately assess the situation
Incorporate Plan or Strategy Conflicts	Challenges responders to de-conflict situation assessment and monitoring plans and strategies

Relevant References

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, *37*(1), 32–64.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A Temporally Based Framework and Taxonomy of Team Processes. *The Academy of Management Review*, *26*(3), 356–376. http://doi.org/10.2307/259182

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

Target Capabilities List: A companion to the National Preparedness Guidelines. (2007, September). U.S. Department of Homeland Security.

8.2.3.1 Assess Technical Hazards (Sub-Capability)

Description
The ability to assess the technical hazard being faced as a result of the spill. This includes the ability to assess if and to what extent oil discharge is active, whether or not the vessel or rig is stable, and if fire suppression is required.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders appropriately analyzed hazard information and consulted experts to assess the active discharge of oil, stability of vessel or rig and if fire suppression was required.	**Result:** Yes or No
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 2 (Output)	
Responders accurately assessed the active discharge of oil, stability of vessel or rig and if fire suppression was required.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

8.2.3.2 Project Future Situations, Issues, and Concerns (Sub-Capability)

Description

The ability to consider future situations, issues, and concerns as it relates to current decisions and actions. This includes considering future meetings, hazards, resources, information, etc. that may impact operations in the future. For example, a decision maker may consider the future arrival time of certain resources and decide that different resources can be used in the meantime.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders considered possible future situations, issues, and concerns in current decision making.	**Result:** 1. Never 2. Somewhat Often 3. Often
Data collection methods: • Observation	**Comments:**

Measure 2 (Output)	
Responders provided clear and defensible projections on a consistent basis.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

8.2.4　Utilize Management Structure with Operating Norms

Description

Ability to utilize and operate within a management structure capable of mobilizing people and equipment as needed for the response. Management structures must scale to responses that are small or large, complex or simple, and may interface with other stakeholders and organizations. In addition to utilizing a suitable management structure for the response, responders should have and use processes and tools to help facilitate response operations (e.g., resource requests, legal review, meetings, operational periods, software systems, etc.).

Evaluation Measures and Techniques

Measure 1 (Output)	
Responders developed and/or utilized a management structure that was clear and adequate to the scale and complexity of the response.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders clearly defined and/or utilized processes that supported the organized response (e.g., processing resource requests).	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 3 (Outcome)	
Responders effectively organized themselves such that they were able to scale the response, assign/delegate responsibilities, and maintain compliance with laws and regulations.	**Result:** 1.　Not Successful 2.　Marginally Successful 3.　Successful 4.　Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Updates, Reports or Information	EX: Require responders to develop a report that outlines the organizational structure or operating norms
Incorporate Ambiguous Group Structures and Operating Norms	Challenges responders develop organizational structure and operating norms
Require Meetings	Prompts responders to discuss organizational structure and operating norms
Incorporate Incomplete or Conflicting Information	Challenges responders to develop organizational structure and operating norms with incomplete or conflicting information
Require Plans or Strategies	EX: Require responders to develop an organization or operating plan that outlines structures and processes
Incorporate Changes in the Situation	Challenges responders to have flexible structure and processes or modify them for situation
Require Strategic, Operational or Tactical Decisions	Prompts responders to make decisions about organizational structure and operating norms
Require Authorizations	Challenges responders to create management structure and processes that reflect required authorizations
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Challenges responders to create management structure and processes that reflect response priorities
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Challenges responders to create management structure and processes that enable coordination
Moderate Relevance	
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Challenges responders to consider stakeholders' priorities and goals when creating management structure and processes
Incorporate Multiple Events	Challenges responders to have flexible structure and processes or modify them for situation
Incorporate Unexpected Situations	Challenges responders to have flexible structure and processes or modify them for situation
Require Execution of Plans, Procedures or Processes	Prompts responders to create management structure and operating norms suitable for executing plans, procedures or processes
Incorporate Resource Insufficiencies	Challenges responders to have flexible structure and operating norms to manage situation despite lack of needed resources
Require Group Decision Making	Prompts responders to discuss and make decisions as a group about management structure and operating norms
Weak Relevance	
Incorporate Public Perception of Priorities, Goals or Tasks	Challenges responders to consider public's perception of response priorities, goals or tasks when creating management structure and processes
Incorporate Interpersonal Conflicts	Challenges responders to have organizational structure or processes to effectively deal with interpersonal conflict
Incorporate Events or Tasks that Stress Plans	Challenges responders to have flexible structure and processes or modify them in absence of adequate prior plans
Require Deployment and Demonstration of Equipment	Prompts responders to create management structure and operating norms suitable deploying and demonstrating the use of specific equipment
Incorporate Time Pressure	Prompts responders to create suitable management structure and operating norms with the pressure of time

Relevant References

Bigley, G. A., & Roberts, K. H. (2001). The incident command system: High-reliability organizing for complex and volatile task environments. *Academy of Management Journal*, *44*(6), 1281–1299.

Pearson, C. M., & Mitroff, I. I. (1993). From crisis prone to crisis prepared-a framework for crisis management.pdf. *Academy of Management Executive*, *7*(1), 48–59.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

8.2.4.1 Assign and Delegate Responsibilities (Sub-Capability)

Description

The ability to assign or delegate responsibilities to others such that division of labor is manageable and the responsibilities are appropriate for the person given their knowledge, skills and abilities.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders assigned or delegated roles and responsibilities to appropriate personnel.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders understood their roles and responsibilities within the response.	**Result:** Yes or No
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

8.2.4.2 Scale Operations (Sub-Capability)

Description

The ability to scale operations to match the needs of response without compromising the performance of other capabilities. This is crucial as a response ramps up and the organizational structure expands to meet the needs of the disaster. Operations often become more complex and response organizations should be able to build a response structure without compromising effectiveness. Conversely, this may include scaling down operations when appropriate.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders requested the appropriate personnel and resources to scale operations to meet the needs of the response.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders were able to scale operations to meet the needs of the response.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

8.2.4.3 Maintain and Manage Compliance with Laws and Regulations (Sub-Capability)

Description

The ability to manage and maintain compliance with existing laws and regulations regarding the response and cleanup of the oil spill. When appropriate, responders may seek out exceptions from the appropriate authorities. For example, responders need to ensure they are using appropriate in-situ burn techniques and obtain the required approvals for their use. Or responders need to ensure they operate in accordance with relevant employment laws with hiring local people to assist in cleanup efforts.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders documented their compliance with relevant laws and regulations affecting operations.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
If an authorization or an exception was required for a particular decision or action, responders took the appropriate steps to document and request the authorization or exception.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

8.2.5 Interface with and Manage Stakeholders

Description
Ability to interface between a responsible party's own response management structure (e.g., Spill Management Team) and other stakeholders affected by the spill itself or response activities, including the affected public. If stakeholders such as the USCG and USFW become part of the response management structure, which is typical as a response scales up, they are considered a stakeholder only when they represent their home organizations' interests in any discussions that take place. Decision-making, information sharing, and other critical functions should run smoothly across the interface. Operational and strategic levels often require interfaces.

A critical component of this is to identify and document all the stakeholders to ensure their goals and objectives are understood and integrated (as appropriate) within the response. This needs to happen early on in a response to ensure all appropriate personnel are "in the room" when decisions are made. Any conflicts between organizations' goals and objectives should be identified and managed. Responders should make stakeholders feel included, valued, and updated on the situation in an easily understandable way that does not prompt fear and rash decisions that could throw off the response efforts.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders had a plan or strategy for interfacing with and managing stakeholders.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders considered stakeholders' priorities and goals in a timely manner and used an appropriate method of communication (e.g., discussion, report, email, phone, etc.).	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 3 (Outcome)	
During the response, stakeholders felt their priorities and goals were acknowledged and considered by responders.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Updates, Reports or Information	EX: Require responders to develop a report for key stakeholders
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Challenges responders to consider stakeholder's priorities, goals or tasks when interacting with them
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Challenges responders to coordinate response efforts with stakeholders
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Challenges responders prioritize response efforts given stakeholder' priorities and goals
Require Authorizations	Prompts responders to seek and manage approval process that involve stakeholders
Require Group Decision Making	Prompts responders to discuss and make decisions with stakeholders
Require Strategic, Operational or Tactical Decisions	Prompts responders to make decisions that involve or impact stakeholders
Incorporate Public Perception of Priorities, Goals or Tasks	Challenges responders to consider the public's perception of response priorities, goals or tasks in response efforts
Incorporate Incomplete or Conflicting Information	Challenges responders to consider stakeholders' priorities and goals despite having incomplete or conflicting about the response or their needs
Incorporate Resource Insufficiencies	Challenges responders to effectively interface with and management stakeholders despite a lack of resources (e.g., no personnel to act as liaison)
Require Execution of Plans, Procedures or Processes	Prompts responders to execute a plan, procedure or process for interfacing with and managing stakeholders
Require Plans or Strategies	EX: Require responders to develop plan or strategy to interface with and manage stakeholders
Require Meetings	Prompts responders to interface with and manage stakeholders
Moderate Relevance	
Incorporate Ambiguous "Correct" Decisions	Challenges responders to interface with and manage stakeholders when a decision needs to be made for which there is no clear answer
Incorporate Ambiguous Group Structures and Operating Norms	Challenges responders to effectively interface with and manage stakeholders despite ambiguous organizational structure or operating norms
Incorporate Unexpected Situations	Challenges responders to effectively interface with and manage stakeholders when an unexpected situation occurs
Incorporate Changes in the Situation	Challenges responders to effectively interface with and

	manage stakeholders when the situation changes
Incorporate Unclear Situations or Problems	Challenges responders to effectively interface with and manage stakeholders when there is an unclear situation or problem
Incorporate Multiple Events	Challenges responders to effectively interface with and manage all relevant stakeholders when multiple events are occurring
Incorporate Interpersonal Conflicts	Challenges effective group interactions with stakeholders when there is interpersonal conflict between or among them
Incorporate Plan or Strategy Conflicts	Challenges responders to de-conflict plans or strategies that affect stakeholder (e.g., 2 GRPs with A priority areas)
Weak Relevance	
Incorporate Team-Based Decisions or Actions	Prompts responders to make decisions or take actions about the response with stakeholders

Relevant References

National Response Framework. (2013, May). U.S. Department of Homeland Security.

U.S. Coast Guard Incident Management Handbook. (2001). US Coast Guard.

8.2.6   Coordinate Operations

Description

The ability to coordinate and align interdependent actions during a response toward the established goals and objectives. This includes the ability to coordinate efforts within and across different teams and organizations, which may require responders to surface stakeholder expectations and priorities early in a response. There is a complementary relationship between this capability and the capabilities listed under the Execution Capabilities category because they are dependent on effective coordination. Additionally, this capability differs from Conduct Strategy and Response Planning in that it focuses on aligning current resources and actions so the Execution Capability can be performed. It focuses on execution rather than planning the execution.

Evaluation Measures and Techniques

Measure 1 (Process)	
Instances of coordination problems, such as insufficient or duplicative resources, confusion, conflict, or duplication of effort, were observed during the response.	**Result:** # of coordination problems (Indicator of coordination effectiveness) *Higher number may be due to miscommunication, poor information management, confusion, not properly assigning/delegating responsibility
Data collection methods: • Observation	**Comments:**

Measure 2 (Output/Outcome)	
Responders were able to align interdependent response actions and positively impact response efforts.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Need to coordinate prompts coordination
Incorporate Incomplete or Conflicting Information	Lack of precise information requires coordination for clarity
Incorporate Resource Insufficiencies	Resource insufficiency requires better coordination of limited resources
Require Strategic, Operational or Tactical Decisions	Decisions require coordination to execute
Moderate Relevance	
Incorporate Unexpected Situations	Unexpected situation requires coordination to address.
Require Authorizations	Need for decision or action approval prompts coordination
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Responders need to coordinate with stakeholders to identify and consider their needs
Incorporate Changes in the Situation	A change in the situation requires coordination to address
Weak Relevance	
Incorporate Events or Tasks that Vary in Size and Require Different Response Levels	Size or complexity of scenario prompts coordination
Require Execution of Plans, Procedures or Processes	Plan requires coordination to execute
Require Plans or Strategies	Plan requires coordination to develop
Incorporate Ambiguous "Correct" Decisions	Decision requires coordination to determine best approach

Relevant References

DeShon, R. P., Kozlowski, S. W. J., Schmidt, A. M., Milner, K. R., & Wiechmann, D. (2004). A Multiple-Goal, Multilevel Model of Feedback Effects on the Regulation of Individual and Team Performance. *Journal of Applied Psychology*, *89*(6), 1035–1056. http://doi.org/10.1037/0021-9010.89.6.1035

Kline, T., & McGrath, J.-L. (1998). Development and validation of five criteria for evaluating team performance. *Organization Development Journal*, *16*(3), 19–27.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A Temporally Based Framework and Taxonomy of Team Processes. *The Academy of Management Review*, *26*(3), 356–376. http://doi.org/10.2307/259182

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

Target Capabilities List: A companion to the National Preparedness Guidelines. (2007, September). U.S. Department of Homeland Security.

8.2.7   Provide and Manage Operational Communications

Description
The ability to ensure adequate communications are provided to and managed for responders. This includes radios, Internet access, phones, network infrastructure, software, hardware, etc. to support the other capabilities listed in this framework. Additionally, the appropriate communication policies and protocols should be established to ensure information is communicated by authorized sources. The communications selected by the response team must be appropriate to the task and activities being performed.

Evaluation Measures and Techniques

Measure 1 (Process/Output)	
Responders provided and managed the appropriate communications tools and resources to support response efforts and stakeholders.	**Result:** Yes or No
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 2 (Output)	
Responder developed a suitable communications management protocol or plan to support response efforts and stakeholders.	**Result:** Yes or No
Data collection methods: • Observation • Documentation • Controller/Evaluator Debriefing	**Comments:**

Measure 3 (Outcome)	
Responders were able to effectively communicate with each other during the response using the available communications tools and resources.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Deployment and Demonstration of Equipment	Demonstrate performance of technical communications skills (e.g., network setup/maintenance, use of radio repeaters, etc.)
Moderate Relevance	
Require Plans or Strategies	Require a communications plan/strategy or system management protocol/plan
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Need to coordinate prompts need for communications
Incorporate Unclear Situations or Problems	Unclear problem requires communications to investigate/troubleshoot
Incorporate Resource Insufficiencies	Resource insufficiency requires communications to coordinate workaround.
Incorporate Ambiguous "Correct" Decisions	Ambiguous "correct" decision requires communications to make final decision
Require Strategic, Operational or Tactical Decisions	Decisions require communications to coordinate, discuss and retrieve information
Require Execution of Plans, Procedures or Processes	Require execution of communications plan or strategy
Weak Relevance	
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Responders require communications with others to coordinate with stakeholders to identify and consider their needs
Incorporate Public Perception of Priorities, Goals or Tasks	Responders require communications to coordinate with public to address their perceptions and resolve or satisfy their needs
Incorporate Changes in the Situation	A change in the situation requires communications to address and coordinate response activities
Incorporate Unexpected Situations	An unexpected situation requires communications to address and coordinate response activities
Require Updates, Reports, or Information	Require a status report on the deployed communications equipment and infrastructure
Incorporate Incomplete or Conflicting Information	Incomplete or conflicting information requires communications to investigate/troubleshoot

Relevant References

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

Target Capabilities List: A companion to the National Preparedness Guidelines. (2007, September). U.S. Department of Homeland Security.

8.2.8   Manage and Share Information

Description

Manage and share information in a timely manner so as to provide relevant information as required and reduce duplication of effort, redundancy and confusion. This capability focuses on the effective acquisition, management, and sharing of information to support responders in their respective roles. For example, available information that does not reach a responder in need of it may be considered poor management and sharing of information.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders had a plan or established process to provide available, timely and relevant information in an appropriate format to other responders (e.g., via document or software).	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Information was effectively managed such that it reduced duplication of effort, redundancy and confusion.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 3 (Outcome)	
Decision makers successfully collected, managed, analyzed and shared relevant information for their decisions and actions using a minimum amount of time and effort.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Updates, Reports or Information	EX: Require responders to develop a report with key information about the response
Incorporate Incomplete or Conflicting Information	Challenges responders to manage and share information when it is incomplete or conflicting
Require Strategic, Operational or Tactical Decisions	Prompts responders to access information to support decisions
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Challenges responders to find and access key information that helps prioritize response efforts
Require Plans or Strategies	EX: Require responders to develop plan or strategy to manage and share information
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Challenges responders to effectively management and share information to support response coordination
Incorporate Changes in the Situation	Challenges responders to effectively manage and share information when the situation changes
Require Meetings	Prompts responders to discuss how they will manage or share information
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Challenges responders to consider stakeholders' priorities and goals when managing and sharing information
Require Authorizations	Prompts responders to seek and manage approval process managing and sharing information
Require Execution of Plans, Procedures or Processes	Prompts responders to create a plan, procedure or process for managing and sharing information if none exist
Require Group Decision Making	Prompts responders to discuss and make decisions as a group about the management and sharing of information
Incorporate Unexpected Situations	Challenges responders to effectively manage and share information when an unexpected situation occurs
Moderate Relevance	
Incorporate Ambiguous Group Structures and Operating Norms	Challenges responders to effectively manage and share information despite ambiguous organizational structure or operating norms
Incorporate Resource Insufficiencies	Challenges responders to effectively manage and share information despite a lack of needed information tools (e.g., incident management software)
Incorporate Public Perception of Priorities, Goals or Tasks	Challenges responders to consider the public's perception of response priorities, goals or tasks when managing and sharing information
Incorporate Events or Tasks that Stress Plans	Challenges responders to effectively manage and share information for response in absence of adequate prior plans
Incorporate Multiple Events	Challenges responders to effectively manage and share information when multiple events are occurring
Incorporate Unclear Situations or Problems	Challenges responders to effectively manage and share information when there is an unclear situation or problem requiring information to address
Weak Relevance	
Incorporate Interpersonal Conflicts	Challenges ability to manage or share information (e.g., conflict makes someone not want to share key information)
Incorporate Time Pressure	Challenges responders to manage and share information with the pressure of time
Incorporate Ambiguous "Correct" Decisions	Challenges responders to manage and share all information needed to support a decision for which there is no clear answer

Incorporate Plan or Strategy Conflicts	Challenges responders to de-conflict plans or strategies (e.g., 2 GRPs with A priority areas)
Incorporate Team-Based Decisions or Actions	Prompts responders to make decisions or take actions as a group about the management and sharing of information
Incorporate Time Flow Adjustments	Prompts responders to manage and share information at different times within a response when things may be more complex

Relevant References

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

Target Capabilities List: A companion to the National Preparedness Guidelines. (2007, September). U.S. Department of Homeland Security.

8.2.8.1 Retain Data and Information (Sub-Capability)

Description

The ability to catalogue, organize and retain documents, information, and data during the response for future litigation, research, or learning. A formal strategy may be developed to help collect and retain material such as ICS forms, emails, notes, databases, claims, etc.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders had a plan or process for cataloguing, organizing, and retaining response data and information for compliance, legal discovery, and learning/evaluation.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders successfully catalogued, organized, and retained relevant response data and information for compliance, legal discovery, and learning/evaluation.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

8.2.8.2 Manage Continuous Updates (Sub-Capability)

Description

The ability to update information on documents as well as information given to others on a continuous basis. Information such as update time, date and audience should be included to make clear what is most recent. This capability helps to avoid operational miscommunications by ensuring everyone has the most up-to-date information related to their roles. This helps everyone maintain a "current" understanding of the situation and issues.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders had a plan or process for continuously updating other responders and stakeholders with information relevant to their operations or priorities.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders and stakeholders were provided continuous and timely updates that were relevant to their operations or priorities and that clearly indicated the most recent update.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

8.2.8.3 Create and Maintain Shared Situation Awareness (Sub-Capability)

Description

The ability to create and document a shared understanding of the situation among team or organizational members. This includes current and possible future issues being faced, the response priorities, and the approach(es) being used to respond to the situation. Creating shared situation awareness may occur in various ways such as through briefings, meetings, software, documents, display screens, phone conversations, etc. These activities are easier to observe for evaluation purposes, but only provide evidence that shared situation awareness activities are taking place, not that good situation awareness has been achieved.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders created and utilized tools for shared situation awareness such as plans, meeting, reports, common operating pictures, etc.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders regularly updated and maintained tools such as plans, meetings, reports, and common operating pictures for shared situation awareness in a way that is easy to use and understand.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant References

Bigley, G. A., & Roberts, K. H. (2001). The incident command system: High-reliability organizing for complex and volatile task environments. *Academy of Management Journal*, *44*(6), 1281–1299.

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, *37*(1), 32–64.

Harrald, J. R. (2006). Agility and Discipline: Critical Success Factors for Disaster Response. *The Annals of the American Academy of Political and Social Science*, *604*, 256–272. http://doi.org/10.1177/0002716205285404

8.2.8.4 Utilize Public Information and Intelligence (Sub-Capability)

Description

The ability to identify, manage and share information provided by the public, which includes ensuring that this relevant information is sent to the appropriate people in a timely and appropriate manner. The public can often provide valuable information to responders during an oil spill response. For example, social media may provide key information from boaters on where oil spill slicks are spreading.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders had a plan or strategy for utilizing public information and intelligence within operations (e.g., social media data or information, a reporting hotline, etc.).	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Responders incorporated public information and intelligence into operations, including decision making, coordination and discussions, etc.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

8.2.9 Manage and Account for Resources

Description

The ability to assess requirements, acquire, account for and manage the flow of human, financial and material resources within the organization's authority or control. This includes establishing the appropriate management structure for all sizes of a disaster and enabling the following resource management activities:

1. Source and Procure Resources
2. Mobilize Response Resources
3. Track and Report Personnel, Assets, and Finances
4. Maintain and Service Equipment
5. Support Response Personnel

Evaluation Measures and Techniques

Measure 1 (Process/Output)	
Responders set up and utilized an effective system for managing and tracking resources (e.g., personnel, assets, and finances).	**Result:** Yes or No
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 2 (Outcome)	
Resource acquisition or management problems did not hinder the oil spill response.	**Result:** 1. Serious Hindrance 2. Slight Hindrance 3. Limited or No Hindrance 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Prompts need to manage and account for dispersed and diverse resources
Require Strategic, Operational or Tactical Decisions	EX: Where should the available resources be assigned? What resources should I procure?
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Prompts need to assign resources to priority areas
Incorporate Resource Insufficiencies	EX: What other resources are available to address issue?
Incorporate Incomplete or Conflicting Information	Challenges ability to account for resources (e.g., unclear if Boat X is available or assigned to task)

Incorporate Changes in the Situation	Challenges ability to manage resources (e.g., reassign resources)
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	EX: 45 fishing vessels are available to support cleanup operations, (e.g., must manage and account for them in response)
Incorporate Unexpected Situations	Challenges ability to manage resources (e.g., 5 boats are disabled due to surface oil caught in engine)
Require Deployment and Demonstration of Equipment	Prompts need to physically deploy and manage specific resources
Moderate Relevance	
Require Authorizations	EX: Dispersant production requires approval to initiate
Require Updates, Reports or Information	EX: Require a report detailing the status and position of every resource (including personnel, assets and finances)
Require Plans or Strategies	EX: Require resource management plan that describes how personnel, asset, and financial resources will be managed and accounted for
Incorporate Multiple Events	Challenges ability to manage resources for multiple events
Require Meetings	EX: Require resource management meeting to discuss how resources will be managed and accounted for
Weak Relevance	
Incorporate Public Perception of Priorities, Goals or Tasks	EX: The perception is the response is being poorly managed because there are many boats in port instead on the water
Require Execution of Plans, Procedures or Processes	EX: Require execution of resource management plan to manage and account for resources
Incorporate Ambiguous "Correct" Decisions	EX: There are multiple impact areas, where should the resources be assigned first?

Relevant References

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

National Incident Management System. (2008, December). U.S. Department of Homeland Security.

National Preparedness for Response Exercise Guidelines (DRAFT). (2015, March). USCG, EPA & BSEE.

National Preparedness for Response Exercise Program Guidelines. (2002, August). USCG, EPA, & DOI.

National Preparedness Goal. (2015, September). U.S. Department of Homeland Security.

National Response Framework. (2013, May). U.S. Department of Homeland Security.

NIMS Resource Management – IS-703.A Course. (2010, January). FEMA Emergency Management Institute.

8.2.9.1 Source and Procure Resources (Sub-Capability)

Description

The ability to assess requirements, locate, list and procure various personnel, asset, and financial resources for the response. This capability focuses on understanding which organizations possess resources helpful to a response, including having insight to the different types of available resources, response times, resource capabilities, etc. Additionally, this capability focuses on having the appropriate resources and processes in place to procure and pay for the resources.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders utilized pre-existing resource contracts and purchase agreements to identify and source needed resources.	**Result:** Yes or No
Data collection methods: • Observation • Participant Hotwash	**Comments:**

Measure 2 (Output)	
Responders were able to quickly identify specific resources that are needed and available for the ***current*** response (e.g., actual inventories or resources that can be procured for the response if it were a real-world incident).	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

8.2.9.2 Mobilize Response Resources (Sub-Capability)

Description

Identify and mobilize resources sufficient to respond to the developing emergency scenario. Resources should be (1) appropriate to the requirements and (2) have sufficient capacity.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders encountered problems with identifying and mobilizing resources needed for the response, such insufficient quantities of resources, wrong resources mobilized, or resources that were not available or mobilized according to contract.	**Result:** \# of problems with resource requests
Data collection methods: • Observation	**Comments:**

Measure 2 (Output)	
Responders were able to quickly identify and mobilize the resources needed to respond to the oil spill, including staging or obtaining operational control of the resources.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

8.2.9.3 Track and Report Personnel, Assets, and Finances (Sub-Capability)

Description

The ability to track and report on the location, status and details of all personnel, assets, and financial resources in order to maintain a clear understanding of all the resources being used in the response. This capability helps with operational coordination and helps improve resource accountability.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders were able to track, with limited confusion and duplication of effort, the status of personnel, assets and finances being used in the response.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 2 (Output)	
Responders provided useful reports on the status of personnel, assets, and finances to other responders.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

8.2.9.4 Maintain and Service Equipment (Sub-Capability)

Description

The ability to monitor, maintain and service response equipment during a response, including the ability to repair equipment and provide supporting services such as power and appropriate support vessels and personnel with minimal disruption to operations.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders developed and utilized a maintenance and service plan for deployed equipment.	**Result:** Yes or No
Data collection methods: • Observation • Documentation	**Comments:**

Measure 2 (Output)	
Equipment maintenance and service problems were effectively addressed in a timely manner.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

8.2.9.5 Support Response Personnel (Sub-Capability)

Description

The ability to provide necessary lodging, food, transportation, healthcare and equipment to response personnel during a response.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders utilized pre-existing contracts and agreements to provide lodging, food, transportation, healthcare and equipment for response personnel.	**Result:** Yes or No
Data collection methods: • Observation • Participant Hotwash	**Comments:**

Measure 2 (Output)	
Responders were able to provide field personnel with adequate lodging, food, transportation, healthcare, and equipment they need in order to respond.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

8.3 Skills-Based Capabilities

8.3.1 Deploy and Manage Effective Teams

Description

The ability to deploy and manage effective response teams at different levels of the response (e.g., operations center vs. boat crew). Teams should have the ability to identify, coordinate and build trust with other teams or stakeholders as well as operate under good leaders and managers that:

- effectively manage conflict and personalities
- build confidence and motivate team members
- provide ongoing and timely feedback to team members
- build trust and relationships within the team

Evaluation Measures and Techniques

Measure 1 (Process)	
Team leaders and managers exhibited behaviors consistent with effective leadership and management. (e.g., building confidence, motivating, providing feedback, building trust, health conflict, etc.)	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 2 (Outcome)	
Team leaders and managers successfully led and managed a team under stress and pressure whereby a healthy culture of conflict, trust, safety confidence and motivation were built.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Requires leadership to manage stakeholders and their needs
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Need to coordinate prompts a need to lead a team

Incorporate Interpersonal Conflicts	Interpersonal conflict prompts leadership action
Require Meetings	A meeting requires leadership and management to execute
Incorporate Public Perception of Priorities, Goals or Tasks	Requires leadership to manage public perception and influence on the response
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Need to prioritize prompts the need for leadership as well as need for leadership to select appropriate course of action
Incorporate Events or Tasks that Exceed Experience or Expertise of Responders	Stress challenges the ability to provide good leadership and management
Moderate Relevance	
Require Plans or Strategies	Plan requires leadership and management to develop
Incorporate Ambiguous "Correct" Decisions	Decision requires leadership to make best decision
Incorporate Changes in the Situation	Requires leadership and management to adapt
Incorporate Incomplete or Conflicting Information	Challenges leadership to make decisions or take action from incomplete or conflicting information
Require Authorizations	Need for decision or action approval requires management of the approval process
Incorporate Unexpected Situations	Requires leadership and management to adapt
Weak Relevance	
Require Strategic, Operational or Tactical Decisions	Good leadership and management is required to make good decisions
Require Updates, Reports or Information	Responders need to work with others to compile the report
Incorporate Unclear Situations or Problems	Requires good leadership to investigate, identify and define the root of the problem
Incorporate Multiple Events	Requires leadership to prioritize efforts and manage additional issues.
Incorporate Events or Tasks that Stress Plans	Challenges leadership to deal with and manage new or overwhelming issues not previously planned for
Incorporate Events or Tasks that Vary in Size and Require Different Response Levels	Challenges leadership to deal with more complex situations
Incorporate Resource Insufficiencies	Requires leadership and management to adapt response
Require Group Decision Making	Requires leadership to set process for decision making
Incorporate Time Flow Adjustments	Challenges leadership to deal with a series of complex situations that would not occur normally in succession

Relevant References

Cantu, C. J. (2007). Evaluating team effectiveness: examination of the team assessment tool. University of North Texas.

Kline, T., & McGrath, J.-L. (1998). Development and validation of five criteria for evaluating team performance. *Organization Development Journal*, *16*(3), 19–27.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A Temporally Based Framework and Taxonomy of Team Processes. *The Academy of Management Review*, *26*(3), 356–376. http://doi.org/10.2307/259182

8.3.2   Solve Problems and Make Decisions

Description

The ability of a team or organization to effectively solve problems and make decisions. The team or organization should have a process that supports individual and/or collaborative problem solving and decision making. Good processes incorporate the usage of available information, consider multiple alternatives, seek expertise and counsel when appropriate (including alternative viewpoints), and coordinate and integrate decisions with others as required. In addition, good processes recognize that the situation is evolving and that there are uncertainties in the information available, the decisions being made, and possible future conditions.

Evaluation Measures and Techniques

Measure 1 (Process)	
The processes used to facilitate problem solving and decision making included the following elements, as needed: usage of available information; consideration of multiple alternatives; outside or alternative expertise, viewpoints and counsel; and decisions or actions that are coordinated and integrated with others.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 2 (Outcome)	
Responders *successfully* solved problems or made key decisions using a defined process.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Require Updates, Reports or Information	EX: Require responders to develop a report on outstanding problems or key decisions to address
Require Strategic, Operational or Tactical Decisions	Prompts responders to make decisions that involve or impact stakeholders
Incorporate Incomplete or Conflicting Information	Challenges responders to solve problems and make decisions despite having incomplete or conflicting information
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Challenges responders to solve problems or make decisions about what response efforts to prioritize
Require Plans or Strategies	EX: Require responders to develop plan or strategy to solve specific problems or make specific decisions
Incorporate Changes in the Situation	Challenges responders to solve new problems and make new decisions when the situation changes
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Challenges responders to solve problems and make decisions that arise when response efforts need to be coordinated
Require Authorizations	Prompts responders to address specific problems or decisions that would affect the authorization process for other decisions or actions
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Challenges responders to solve problems and make decisions that involve and impact stakeholders
Require Group Decision Making	Prompts responders to discuss problems and make decisions as a group
Incorporate Resource Insufficiencies	Challenges responders to solve problems and make decisions when confronted with a resource insufficiency
Incorporate Unexpected Situations	Challenges responders to solve new problems and make new decisions when an unexpected situation occurs
Incorporate Ambiguous "Correct" Decisions	Challenges responders to make decisions for which there is no clear answer
Incorporate Plan or Strategy Conflicts	Challenges responders to de-conflict plans or strategies (e.g., 2 GRPs with A priority areas)
Moderate Relevance	
Incorporate Public Perception of Priorities, Goals or Tasks	Challenges responders to consider the public's perception of response priorities, goals or tasks when solving problems and making decisions
Require Meetings	Prompts responders to discuss problems and decisions as a group
Incorporate Multiple Events	Challenges responders to solve problems and make decisions when multiple events are occurring
Incorporate Ambiguous Group Structures and Operating Norms	Challenges responders to solve problems and make decisions despite an ambiguous organizational structure or operating norms
Incorporate Interpersonal Conflicts	Challenges effective problem solving and decision making when there is interpersonal conflict between or among them a group
Incorporate Events or Tasks that Stress Plans	Challenges responders to solve problems and make decisions in absence of adequate supporting plans
Weak Relevance	
Require Execution of Plans, Procedures or Processes	Prompts responders to execute a plan, procedure or process for solving problems and making decisions
Require Deployment and Demonstration of Equipment	Prompts responders to solve problems and make decisions that arise when equipment is deployed and demonstrated (e.g., equipment vessel with too large of draft for area)

Relevant References

Eisenhardt, K. M. (1989). Making fast strategic decisions in high-velocity environments. *Academy of Management Journal*, *32*(3), 543–576.

Kline, T., & McGrath, J.-L. (1998). Development and validation of five criteria for evaluating team performance. *Organization Development Journal*, *16*(3), 19–27.

8.3.2.1 Evaluate Alternatives (Sub-Capability)

Description

The ability to consider multiple options for response without becoming hyper-focused on a particular decision or action. Evaluation of alternatives may occur implicitly or explicitly in discussions, but should be thorough enough for the given situation and documented well-enough to explain decision rationale. For example, when considering where to obtain chemical dispersant sufficient for a large oil spill, a decision maker will give ample consideration to the benefits/ negatives of purchasing from a vendor or asking the manufacturer to ramp up production. The decision maker may consider lead times as well as the chemical properties of the available dispersant to develop a strategy or make a decision.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders adequately documented and explained their decision rationale.	**Result:** Yes or No
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Measure 2 (Output)	
Responders solved a problem or made a decision in which multiple alternatives were compared and evaluated in an impartial manner or without bias toward a particular solution.	**Result:** 1. Did Not Perform Well 2. Performed Satisfactorily 3. Performed Well
Data collection methods: • Observation • Controller/Evaluator Debriefing	**Comments:**

8.3.3 Adapt and Be Flexible

Description

The ability of a team or organization to adapt to the dynamic and evolving situation during response operations. This includes the ability to recognize when adaptation or flexibility is needed. For example, a responder may need to depart from pre-specified procedures due to the situation.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders recognized the need to adapt their plans, procedures or actions when the situation changed from what was expected or planned for.	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 2 (Output)	
Responders only deviated from established plans, procedures or actions when there was clear indication that a different or modified approach may have been warranted.	**Result:** Yes or No
Data collection methods: • Observation • Controller/Evaluator Debriefing	**Comments:**

Measure 3 (Outcome)	
Responders were able to achieve their assigned responsibilities by adapting plans, procedures and actions to the current situation. For *lessons learned*, how were responders able to adapt?	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Documentation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Real-time coordination requires personnel to adapt and be flexible to situation
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Prompts need to adapt and be flexible to deal with response priorities
Incorporate Unclear Situations or Problems	Challenges responders to be flexible in the face of uncertainty
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Challenges responders to address stakeholder needs even though it may not align with current priorities or operations
Incorporate Events or Tasks that Stress Plans	Challenges responders to adapt to current situation rather than planned situation
Incorporate Resource Insufficiencies	Challenges responders to adapt to situation given the resources available
Incorporate Changes in the Situation	Challenges responders to recognize and adapt to changes in situation
Incorporate Plan or Strategy Conflicts	Challenges responders to adapt and be flexible when appropriate course of action is conflicting
Moderate Relevance	
Incorporate Incomplete or Conflicting Information	Challenges responders to be flexible in the face of uncertainty
Incorporate Ambiguous "Correct" Decisions	Challenges responders to adapt and be flexible when appropriate course of action is ambiguous
Weak Relevance	
Require Strategic, Operational and Tactical Decisions	Requires responders to incorporate flexibility into decisions (e.g., in-situ burn testing, if not working, use dispersants)
Require Plans or Strategies	Requires responders to incorporate flexibility into plans
Require Deployment and Demonstration of Equipment	Ex: If primary equipment unavailable, responders demonstrate use of alternative equipment

Relevant References

Bigley, G. A., & Roberts, K. H. (2001). The incident command system: High-reliability organizing for complex and volatile task environments. *Academy of Management Journal*, *44*(6), 1281–1299.

Ford, J. K., & Schmidt, A. M. (2000). Emergency response training: strategies for enhancing real-world performance. *Journal of Hazardous Materials*, *75*, 195–215.

Harrald, J. R. (2006). Agility and Discipline: Critical Success Factors for Disaster Response. *The Annals of the American Academy of Political and Social Science*, *604*, 256–272. http://doi.org/10.1177/0002716205285404

8.3.4   Utilize Adaptive Expertise

Description

The ability to recognize relevant patterns and knowledge from previous disasters or experiences useful to the current situation. This includes the ability to adapt and apply this knowledge despite differences between the prior and current situations. One demonstration of adaptive expertise is the ability to recognize the need for or applicability of a given procedure or strategy to a new situation.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders recognized similar previously experienced situations and their applicability to the current scenario. (You might be able to observe this if responders refer to previous situations as justification for actions or ideas.)	**Result:** \# of instances observed *Note that a higher number isn't necessarily better, but demonstrating this kind of reasoning in a few instances is sufficient.
Data collection methods: • Observation	**Comments:**

Measure 2 (Output)	
Responders were able to appropriately adapt and apply previous knowledge or experience to benefit the current situation.	**Result:** Yes or No
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Prompts responders to recall prior actions/knowledge that might be helpful to coordination
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Prompts responders to recall prior actions/knowledge that might be helpful to prioritizing response efforts
Moderate Relevance	
Incorporate Events or Tasks that Stress Plans	What happened last time something wasn't in a plan?
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	What happened last time when parish wanted more boom now!
Require Strategic, Operational or Tactical Decisions	Why did we decide last time to use ISB, booming, and/or dispersants?
Incorporate Ambiguous "Correct" Decisions	Was there a similar situation faced last time?
Incorporate Incomplete or Conflicting	How was incomplete or conflicting information dealt

Information	with previously?
Incorporate Resource Insufficiencies	We were able to use Boat X last time even though it is not built for it.
Weak Relevance	
Incorporate Changes in the Situation	Prompts responders to recall when a situation changed in a previous times and how it might be applicable now
Require Plans or Strategies	Prompts responders to recall relevant prior knowledge as they are forced develop the plan.
Incorporate Deployment and Demonstration of Equipment	Prompts responders to recall prior actions/knowledge that might be helpful in using equipment
Incorporate Plan or Strategy Conflicts	How was plan/strategy conflict dealt with previously?

Relevant References

Ford, J. K., & Schmidt, A. M. (2000). Emergency response training: strategies for enhancing real-world performance. *Journal of Hazardous Materials*, *75*, 195–215.

8.3.5   Improvise

Description

The ability to creatively respond when existing tools, procedures or operating norms are not developed or not applicable to the situation at hand. Responders, including the response organization, are able to improvise with different tools, rules and, routines that may not have been trained on or planned for in advance. This improvisation enables the accomplishment of goals and objectives despite the unfamiliar conditions.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders accurately identified situations when existing tools, procedures or operating norms were not applicable or were not successful.	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 2 (Output)	
Responders' attempts at improvisation led to an improved response effort.	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 3 (Outcome)	
Responders were able to achieve their goals and objectives despite the lack of appropriate or available tools, procedures, or operating norms. For *lessons learned*, how were responders able to achieve their goals and objectives?	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Resource Insufficiencies	Needed skimming vessels are 3+ days out (e.g., deploys available experimental oil cleanup vessels)
Moderate Relevance	
Incorporate Unexpected Situations	Primary skimming vessel Y broken down. (e.g., use available Boat X even though it wasn't in the plan and is not built for mission)
Incorporate Changes in the Situation	Current begins carrying oil West instead of East as anticipated, but strategy built for East direction
Incorporate Time Pressure	Newly identified hurricane is projected to arrive within three days, creating possible need to evacuate all responders to safe locations in time
Incorporate Plan or Strategy Conflict	Geographic Response Plans conflict with Area Contingency Plan
Specify Precipitating Event	If the precipitating event is unplanned for, improvisation would be required.

Relevant References

Bigley, G. A., & Roberts, K. H. (2001). The incident command system: High-reliability organizing for complex and volatile task environments. *Academy of Management Journal*, *44*(6), 1281–1299.

Ford, J. K., & Schmidt, A. M. (2000). Emergency response training: strategies for enhancing real-world performance. *Journal of Hazardous Materials*, *75*, 195–215.

8.3.6 Prioritize Response Efforts

Description

The ability to prioritize response efforts when there are competing or multiple demands for capabilities and resources. Responders, including the response organizations, are able to identify the most important aspects of the situation and address them with the resources they have available. Prioritizing enables the accomplishment of the most impactful activities given the limited resources and time constraints present in an emergency or disaster.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders recognized the need to prioritize response efforts and established a process and timeline to prioritize efforts.	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 2 (Output)	
Responders were able to prioritize response efforts among competing demands for capabilities and resources.	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 3 (Outcome)	
Responders ability to prioritize response efforts positively impacted response operations.	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Challenges responders to prioritize competing demands or priorities
Require Strategic, Operational or Tactical Decisions	Prompts responders to make specific decisions that prioritize response efforts
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Challenges responders to coordinate response efforts so that priority issues are addressed first
Require Group Decision Making	Prompts responders to make decisions about response efforts as a group
Incorporate Resource Insufficiencies	Challenges responders to re-prioritize efforts if needed resource(s) are unavailable or unable
Incorporate Stakeholder Influence on Priorities, Goals or Tasks	Challenges responders to consider stakeholder's priorities, goals and tasks within larger response
Incorporate Ambiguous "Correct" Decisions	Challenges responders to prioritize response efforts despite no clear and correct decision
Incorporate Public Perception of Priorities, Goals or Tasks	Challenges responders to consider the public's priorities, goals and tasks within larger response
Incorporate Multiple Events	Challenges responders to prioritize responder efforts
Moderate Relevance	
Incorporate Incomplete or Conflicting Information	Challenges responders to prioritize response efforts with incomplete or conflicting information
Require Plans or Strategies	EX: Require responders to develop plan that outlines response priorities.
Require Updates, Reports or Information	EX: Require responders to develop a report that provides an update on response progress and established priorities.
Require Meetings	Prompts responders to discuss response priorities
Require Authorizations	Prompts responders to consider others' response priorities
Incorporate Plan or Strategy Conflicts	Challenges responders to de-conflict plans or strategies (e.g., 2 GRPs with A priority areas, but resources for 1)
Incorporate Changes in the Situation	Challenges responders to re-prioritize response efforts
Incorporate Unexpected Situations	Challenges responders to re-prioritize response efforts
Incorporate Unclear Situations or Problems	Challenges responders to develop priorities despite no clear understanding of the problem
Incorporate Ambiguous Group Structures and Operating Norms	Challenges responders to prioritize response efforts despite not having an established decision making process
Incorporate Team-Based Decisions or Actions	Challenges responders to prioritize response efforts together
Weak Relevance	
Incorporate Deployment and Demonstration of Equipment	Prompts responders to deploy and demonstrate the use of specific equipment

8.3.7   Utilize Prior Plans to Support Operations

Description
The ability to identify and utilize pre-developed documents or plans such as an OSRP, GRPs, maps, etc. that are applicable to the current spill response. Exercise participants should capitalize on prior knowledge without re-inventing prior work or bypassing potentially useful information that is available. Pre-developed documents or plans can be used as is, be updated to meet the specific needs of the response, or be referenced as needed to support operations. This capability is often used in conjunction with other capabilities such as Strategy and Response Planning, Problem Solving and Decision Making and Adaptability and Flexibility.

Evaluation Measures and Techniques

Measure 1 (Process)	
Responders considered or referred to pre-developed documents or plans to inform current activities.	**Result:** Yes or No
Data collection methods: • Observation	**Comments:**

Measure 2 (Outcome)	
The consideration or reference to pre-developed documents or plan positively impacted response operations (e.g., quicker development of plan, more informed decision, etc.).	**Result:** 1. Not Successful 2. Marginally Successful 3. Successful 4. Unknown/Not Demonstrated
Data collection methods: • Observation • Participant Hotwash • Controller/Evaluator Debriefing	**Comments:**

Relevant Exercise Design Components

Exercise Design Component	Explanation/Example
Strong Relevance	
Incorporate Events or Tasks that Require Prioritization of Response Efforts	Prior planning work helps identify operational priorities
Require Strategic, Operational or Tactical Decisions	Prompts responders to access relevant plans and procedures
Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel	Requires usage of prior plans to know how to organize and coordinate operations
Require Plans or Strategies	Prompts responders to access plans and procures that may be relevant to current situation
Incorporate Incomplete or Conflicting Information	Challenges responders to identify and address planning gaps (e.g., plan that only addresses step 1 of 3)
Incorporate Stakeholder Influence on	Challenges responders to know when to deviate from plans

Priorities, Goals or Tasks	and procedures
Moderate Relevance	
Incorporate Resource Insufficiencies	Challenges responders to rethink plans that were dependent on particular resources
Incorporate Changes in the Situation	Challenges responders to rethink plans that were dependent on a particular situation
Require Deployment and Demonstration of Equipment	Prompts responders to identify and use pre-developed plans to deploy and use equipment
Require Authorizations	Challenges responders to ensure plans are up-to-date and applicable to current situation
Incorporate Events or Tasks that Stress Plans	Challenges responders to identify and address planning gaps or assumptions given the current situation
Incorporate Ambiguous "Correct" Decisions	Challenges responders review prior plans that may help inform decisions
Incorporate Plan or Strategy Conflicts	Prompts responders to resolve strategy and planning conflicts (e.g., two GRPs with "A" priorities, but lack of resources for both)
Weak Relevance	
Require Meetings	Ex: Require a search and rescue planning meeting
Require Updates, Reports or Information	Ex: Require an IAP that specifically mentions the progress since the last IAP
Incorporate Unexpected Situations	Challenges responders to consider impact on current plans and if other plans might apply to the situation
Incorporate Public Perception of Priorities, Goals or Tasks	Challenges responders to know when to deviate from plans and procedures
Require Execution Plans, Procedures or Processes	Ex: Require execution of booming strategy in OSRP
Require Group Decision Making	Prompts responders to make decisions together than may be informed by prior plans
Incorporate Conflicting Organizational Goals & Objectives	Challenges responders to know when to deviate from plans and procedures

This page intentionally left blank.

9 APPENDIX B: Exercise Design Components

9.1 Baseline Event

The following three exercise design components are typically part of every exercise. They are designed and implemented in ways that reflect the capabilities being tested and evaluated. Because of this, they have relatively strong links to nearly all capabilities in this framework.

9.1.1   Specify Precipitating Event

Description
Describes the precipitating event that requires a response and its location. In marine oil spills, precipitating events might include pipeline breaks, drilling or well problems, tanker accidents, or even terrorist incidents. Details such as how and where the spill occurred enable exercise participants to understand factors that may complicate containing or stopping the spill and to take actions to mitigate the spill's impact.

9.1.2   Specify Event Location and Impact

Description
Describes the current and anticipated impact of the precipitating event. Details such as the location, type and extent of impact provide implicit information about spill-related risks so they can be assessed and responded to by exercise participants. For example, the impacts might include information such as: 10 boats are contaminated, or the slick of oil is near a common fishing area. With this type of information, exercise participants can begin planning for decontamination and establishing a safety perimeter. The amount and detail of the impacts may be varied; for example, in some cases, anticipated impacts may be described, and in others, the participants are expected to determine anticipated impacts based on available information.

9.1.3   Specify Environmental Conditions

Description
Describes the environmental conditions in which the response takes place, including the weather, wind, ocean current, temperature, and other relevant details. This information challenges participants to determine whether or not they can operate in the area, identify the oil trajectory, and determine required safety precautions, and it influences other major decisions as well. The initial environmental conditions may be provided at the start of the exercise with updates provided as the exercise progresses.

9.2 Baseline Tasks

9.2.1   Require Plans or Strategies

Description
Exercise participants must discuss and/or develop an informal or formal plan or strategy to address particular issues. For example, the incident command team may develop the Incident Action Plan to communicate the intended objectives and activities over the operational period(s). This stimulates exercise participants to clarify and communicate their goals and objectives. Some plans and strategies are already required of participants during the normal course of ICS operations. Exercise designers may choose to require the development of additional plans (e.g., by requesting such a plan within the exercise).

9.2.2   Require Execution of Plans, Procedures or Processes

Description
Participants must execute a plan, procedure or process, either developed prior to the response or during the response. The goal is to determine if the plan can be executed and if any changes are needed for an actual event. For example, a plan may have been developed that is not realistic in practice. Or a plan may have been developed, but the required skills and abilities have not been mastered. This task can help elucidate specific issues with plans, including faulty assumptions, as well as issues with plan execution, such as inability to deploy and coordinate resources.

9.2.3   Require Group Decision Making

Description
Participants must discuss and make decisions as a group. They must demonstrate satisfactory leadership and teamwork skills as well as an adequate decision-making process that includes (1) use of available information, (2) consideration of multiple alternatives, (3) use of specialized expertise and counsel, and (4) integration of decisions with a larger response. This design component enables exercise designers to evaluate group dynamics in a stressed environment. The goal is to understand group dynamics, not to evaluate the efficacy of the decisions. This design component can be implemented by requiring a particular decision be made by a group of several stakeholders, such as a Unified Command.

9.2.4   Require Meetings

Description
Participants must execute a particular type of meeting such as a goals and objectives meeting or a press conference. The goal is to determine if the meeting can be properly executed and to observe group dynamics. This task can help elucidate specific issues including leadership and team dynamics, problem solving and decision making, shared situation awareness, and stakeholder interaction.

9.2.5   Require Updates, Reports or Information

Description

Participants provide a requested report, update or piece of information. For example, they may be asked for details on the situation, the actions being taken, or outstanding resources. The goal is to have exercise participants articulate their understanding of the situation and response, including progress toward goals and objectives. For example, this design component could be implemented by injecting a request for a status report from a stakeholder.

9.2.6   Require Strategic, Operational or Tactical Decisions

Description

Exercise participants must work through a combination of strategic, operational, and tactical decisions. They must demonstrate competence to make decisions at different levels. The goal is to identify their strengths and weaknesses as they relate to specific types of decisions. For example, a person or team may be very good at tactical decision making (e.g., how to deploy containment booms), but may lack operational (e.g., best place to deploy booms) and strategic decision making ability (e.g., number of booms needed in the next two weeks) that would improve outcomes in the long run.

9.2.7   Require Authorizations

Description

Participants must seek approval from other responders or stakeholders in order to finalize and move forward with a decision or plan of action. This requires participants to coordinate and involve others in their decision making and planning efforts.

9.2.8   Require Deployment and Demonstration of Equipment

Exercise participants must successfully deploy and demonstrate the use of technical response equipment. This helps drive more realistic exercise play and enables evaluators to observe key technical capabilities. Without this requirement, it is hard to evaluate whether an operator can perform complex capabilities such as Mitigate Environmental Impact/Contain Spill using boom or dispersant.

9.3 Complexity Factors: Dealing with the Unexpected

9.3.1   Incorporate an Unannounced Exercise or Withheld Scenario

Description

Participants are surprised by an exercise, or they are aware that an exercise is taking place but are not sure of the scenario they will encounter. The goal of this design component is to test readiness to respond quickly to an event or the specifics of an event. This component helps maintain realism because exercise participants are required to respond with the capabilities they have at that moment rather than boost their capabilities specifically for the exercise. This enables exercise designers to assess whether operators have the requisite capabilities to deploy at that time.

9.3.2   Incorporate Resource Insufficiencies

Description

The scenario requires participants to act and adapt to a situation in which the resources they were planning to use are unavailable or inadequate for the situation. Participants must re-think how to deal with the current situation, including how to solve problems and make decisions, improvise, adapt and be flexible, and utilize adaptive expertise.

9.3.3   Incorporate Unexpected Situations

Description

The scenario requires exercise participants to deal with an unexpected event or problem to test participants' abilities to handle the unexpected with and without available procedures. For example, a key piece of equipment or communications fails, requiring exercise participants to problem-solve and decide what to do. This design component can be tested in two ways:
1. *When a contingency procedure is available:* participants must recognize the need for the procedure and execute it.
2. *When a contingency procedure is not available:* participants must recognize there is no procedure and proceed in a reasonable and appropriate manner. In some cases, they may follow heuristic rules in absence of specific procedures. For example, participants may follow a heuristic that outlines the steps they should take when there is a problem with no defined procedure.

Relevant References

Kluge, A., Sauer, J., Burkolter, D., & Ritzmann, S. (2010). Designing Training for Temporal and Adaptive Transfer: A Comparative Evaluation of Three Training Methods for Process Control Tasks. *Journal of Educational Computing Research*, *43*(3), 327–353. http://doi.org/10.2190/EC.43.3.

9.3.4   Incorporate Changes in the Situation

Description
The scenario includes a change or possible change in the emergency situation that the responders are facing. This design component requires participants to adapt to the new or evolving situation without rigidly adhering to previous but no longer effective decisions and strategies. For example, a change in the weather may disperse the oil or delay its containment. This requires a new approach to containing the oil, or a change in goals. Participants need to recognize the situation has changed or will change and take appropriate action.

9.4 Complexity Factors: Dealing with Scale and Time

9.4.1   Incorporate Events or Tasks that Vary in Size and Require Different Response Levels

Description
Events or tasks within the scenario vary in size and tests participants' ability to manage a response at different response levels (e.g., tactical, operational, and strategic). Larger events often present different problems and require different strategies and approaches in order to respond successfully. For example, a small-scale drill or exercise scenario tests interaction on a tactical and task-oriented basis while a larger and more complex exercise scenario tests inter-organizational interaction on a more strategic level, possibly with competing goals and objectives. Alternatively, designers could select other design components to change the response level (e.g. from operational to strategic) without increasing the scenario size.

Relevant References
Harrald, J. R., Marcus, H. S., & Wallace, W. A. (1990). The EXXON Valdez: An Assessment of Crisis Prevention and Management Systems. *Interfaces*, *20*(5), 14–30.

9.4.2   Incorporate Multiple Events

Description
The scenario introduces simultaneous and potentially cascading events in order to test the participants' ability to handle multiple evolving issues, a common occurrence in disasters. For example, an exercise designer may incorporate a ship-to-ship crash with two ships leaking oil simultaneously. Alternatively, a source control issue might become more difficult when four support vessels become trapped in the oil slick, forcing a rescue and reconsideration of resources needed.

9.4.3   Incorporate Time Pressure

Description
The scenario requires participants to operate in a time-constrained environment. This helps determine not just whether the task can be completed, but also whether it can be completed in an appropriate time frame.

Relevant References
Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A Temporally Based Framework and Taxonomy of Team Processes. *The Academy of Management Review, 26*(3), 356–376. http://doi.org/10.2307/259182

9.4.4   Incorporate Time Flow Adjustments

Description
The exercise includes an adjustment in the time flow, such as a jump from day 1 to day 3, to allow (1) participants to encounter complex scenarios that take time to emerge, and (2) the demonstration of capabilities that would normally take hours, days, or weeks to complete. For example, shoreline cleanup may not begin until a week after the spill. To test it, an exercise could require that participants perform the immediate response mobilization on day 0, then jump to day 7 to implement shoreline cleanup. This design component can be implemented in three ways:
1. An exercise may start mid-way through a simulated response. The scenario would reflect actual information available at that time in order for realistic exercise play to begin.
2. A t*ime jump* may be used to fast forward the scenario from one point in time to another. For example, an exercise scenario starts on Day 1 three hours into the incident. The next part of the exercise would start on Day 3 of the response four hours into the operational period.
3. An exercise may also incorporate a *fast clock* to speed up response efforts and add an element of time pressure. This also enables higher level objectives and capabilities such as problem solving and decision making to occur more often in an exercise.

9.5 Complexity Factors: Managing Influences and Conflict

9.5.1   Incorporate Stakeholder Influence on Priorities, Goals or Tasks

Description
The scenario introduces high- and/or low-level stakeholders (from government representatives to a local citizen), who may have differing and conflicting goals that emerge during a response. These goals and objectives may conflict with operational objectives, and exercise participants must manage this conflict in a reasonable and

appropriate manner. For example, a group of fisherman begin demanding immediate compensation for lost revenue.

9.5.2 Incorporate Public Perception of Priorities, Goals or Tasks

Description

The scenario introduces public perception, which may influence or conflict with how exercise participants react to a problem. For example, public perception might be that responders are not addressing a particular region of shoreline for cleanup. This may have been a strategic decision on the part of the responders, but the public feels it is being neglected or that it should be prioritized. Responders may choose to address this issue by developing a new outreach campaign or prioritizing the shoreline in its operations. This design component may also include the need for responders to act preemptively to mitigate bad public perception of an action or situation before public perception exists or is explicitly stated. For example, this design component could be implemented by injecting complaints from public representatives, through challenging questions from the media, or by managing a fake twitter feed.

9.5.3 Incorporate Conflicting Organizational Goals and Objectives

Description

The scenario introduces inter-organizational conflict to test how teams handle competing goals and objectives. For example, environmental agencies want to minimize impact to wildlife and habitats, while local businesses want to protect tourist beaches. The component requires participants to acknowledge and incorporate others' goals and objectives in a reasonable and appropriate way.

9.5.4 Incorporate Events or Tasks that Require Prioritization of Response Efforts

Description

The scenario requires that participants choose what to prioritize among several important actions. For example, resources in a real-world response are often insufficient to accomplish all required actions at once. As a result, participants must develop a meaningful way to prioritize their efforts (e.g., a defensible process for determining priorities) and decide what to prioritize (e.g., prioritize booming because oil is close to shore). This design component could be implemented by creating a resource shortage, or by presenting several stakeholder requirements that cannot all be accomplished.

9.5.5 Incorporate Plan or Strategy Conflicts

Description

The scenario requires usage of plans or strategies that conflict with each other. This requires exercise participants to identify and reconcile their decisions and actions to deal with conflicting directions and to ensure the response is aligned. For example, Geographic Response Plans, Area Contingency Plans, and OSRPs may call for different actions. Which one supersedes? In what instances? Why?

9.6 Complexity Factors: Organizing and Managing People

9.6.1 Incorporate Events or Tasks that Require Coordination Among Dispersed and Diverse Personnel

Description

The scenario requires close coordination among dispersed and diverse personnel in order to successfully respond. Therefore, participants must interact with others and manage the response when responders are not in the same location and/or have differing expertise. This requires participants to operate in a challenging atmosphere that requires excellent communication. This design component could be implemented by ensuring that some team members are not in the same building as the rest of the team; for example, source control could be in a different city, or some team members could be on a ship coordinating the on-water response.

Relevant References

Ford, J. K., & Schmidt, A. M. (2000). Emergency response training: strategies for enhancing real-world performance. *Journal of Hazardous Materials*, *75*, 195–215.

9.6.2 Incorporate Ambiguous Group Structures and Operating Norms

Description

The scenario includes issues for which there are ambiguous or no preexisting group structure and/or operating norms; in other words, participants must develop a team structure and operational approach that facilitates team performance. They must clearly articulate roles and responsibilities as well as the cycle of activities such as meetings and briefings. The structure and norms should be well organized and integrated into the response. This design component tests participants' abilities to develop and adapt organizational structure in an oil response and enables exercise designers to evaluate teamwork and leadership in a dynamic environment. This design component could be implemented by making an emergency larger than the organization was designed to manage, or by creating an emergency that falls outside the established or planned-for organizational structure.

Relevant References

Zantow, K., Knowlton, D., & Sharp, D. (2005). More Than Fun and Games-Reconsidering the Virtues of Strategic Management Simulations.pdf. *Academy of Management Learning & Education*, *4*(4), 451–458.

9.6.3 Incorporate Interpersonal Conflicts

Description

There are often differences in personality that make it difficult to operate as a team. This design component requires participants to effectively manage and work with difficult

personalities and interpersonal conflicts that arise in the course of a stressful response. This helps assess team dynamics and interpersonal leadership. To implement this component in an exercise, a facilitator may need to play the role of the "difficult personality."

9.6.4   Incorporate Team-Based Decisions or Actions

Description
The scenario requires decisions or actions to be accomplished by a team, stimulating participants to build and maintain a productive team environment that supports response efforts. The goal is to create conditions that stress the team in order to identify teamwork characteristics that either positively or negatively affect response efforts. For example, one may want to test the ability to form and operate a geographically dispersed team. Or, one may want to test the ability for a strike team of four different contractors to operate well together despite normally competing with each other.

Relevant References
Gralla, E., Goentzel, J., & Chomilier, B. (2015). Case study of a humanitarian logistics simulation exercise and insights for training design. *Journal of Humanitarian Logistics and Supply Chain Management*, *5*(1), 113–138.

9.7 Complexity Factors: Dealing with Ambiguity

9.7.1   Incorporate Unclear Situations or Problems

Description
The scenario includes an unclear situation or problem, requiring exercise participants to define the problem or problems that exist, identify ways to solve the problem, and then decide on the appropriate actions. The set of information provided during the exercise is purposely ambiguous, so participants do not immediately know what problems they should be solving. The goal is for exercise participants to demonstrate that they can figure out what needs to be done in a complex and evolving response scenario. For example, there is an abnormal surge in the amount of vessels requesting to be decontaminated, which could be for a variety of reasons.

Relevant References
Ekfeldt, B., Österberg, R., Nyström, M., Stamatopoulou, M., Kargakou, E., Konstantarogianni, E., … others. (2015). Preparing for Care in a Combat Environment. *International Journal of Caring Sciences*, *8*(1).

Ford, J. K., & Schmidt, A. M. (2000). Emergency response training: strategies for enhancing real-world performance. *Journal of Hazardous Materials*, *75*, 195–215.

Gralla, E., Goentzel, J., & Chomilier, B. (2015). Case study of a humanitarian logistics simulation exercise and insights for training design. *Journal of Humanitarian Logistics and Supply Chain Management*, *5*(1), 113–138.

Smith, S., Farra, S., Dempsey, A., & Arms, D. (2015). Preparing nursing students for leadership using a disaster-related simulation. *Nurse Educator*. http://doi.org/10.1097/NNE.0000000000000143

9.7.2 Incorporate Incomplete or Conflicting Information

Description

Exercise participants are provided with incomplete or conflicting data or information, which impacts a decision-making process. For example, three weather models are predicting different storm tracks in the next three days. This requires participants to make judgments about next steps despite the uncertainty and determine whether additional information is needed or whether a decision can or must proceed. There are three types of information:

1. *Incomplete* – Not all data or information needed is available or data or information is only representative of a certain area or population.
2. *Conflicting* – Data or information received from different sources are contradictory or outdated.
3. *Expected* – Data or information expected is not provided (e.g., staging team doesn't report the arrival of requested resources).

9.7.3 Incorporate Ambiguous "Correct" Decisions

Description

The scenario requires exercise participants to grapple with a set of options with no clear right or best decision. The goal is to force participants to work through a decision process when there is no clear right answer, then deal with the resulting consequences. For example, the choice to use dispersants, conduct in-situ burning, boom and/or skim is judgement call based on response priorities, available resources, environmental conditions and anticipated impacts.

9.7.4 Incorporate Events or Tasks that Stress Plans

Description

The scenario includes elements that either have not been planned for or for which the plans may be unrealistic or inadequate, aiming to "stress" the plan. Constructing such a scenario may require exercise designers to identify gaps or unrealistic assumptions in existing plans. Examples include exercising a larger scenario than was considered in the plan, in a location outside those mentioned in the plan, with cascading events not considered (together) in the plan, and/or with weather that was not considered in the plan. The goal is to highlight areas in which plans must be improved, and/or to test the organization's adaptability and flexibility when scenarios are unplanned-for.

9.7.5 Incorporate Events or Tasks that Exceed Experience or Expertise of Responders

Description

The scenario requires participants to step outside their comfort zone to deal with complex and challenging situations. The goal is to stress the capabilities of the exercise participants past their existing knowledge, skills, and abilities. This helps mimic real-world conditions in which the type and impact of on oil spill has never been encountered before. It enables exercise designers to evaluate the likelihood that participants can perform well no matter the situation. For example, this design component can be implemented by creating conditions or tasks that are rare and challenging, such as equipment malfunctions that have not been encountered, serious resource insufficiencies, or difficult weather.

Implementing this design component may be problematic under BSEE's current regulatory environment, because if an operator fails an exercise designed with this component, it could be considered the designer's fault rather than the company's fault. We note this issue, but it is important to include this type of exercise design component in this framework, because skills at dealing with unexperienced events may be important in future oil spills, and future regulations may enable them to be tested.

Relevant References
Fletcher, L., Justice, S., & Rohrig, L. (2015). Designing a Disaster. *Journal of Trauma Nursing*, *22*(1), 35–40. http://doi.org/10.1097/JTN.0000000000000098

This page intentionally left blank.

10 APPENDIX C: Exercise Evaluation

The specific evaluation measures for each capability were provided in Appendix A. This appendix contains descriptions of several evaluation techniques that can be used to evaluate each of the measures provided in Appendix A.

10.1 Evaluation Techniques

10.1.1 Observer Ratings

Description
This evaluation technique uses observers to rate participant performance in an exercise. Observers are additional personnel who do not participate in the exercise. Their primary role is to collect data and information to support the evaluation of participant performance (as a result, they fulfill different duties than exercise facilitators or simulators). Observers may be experts in the field or specially trained to evaluate exercises. They look for behaviors and indicators of good performance.

There are a number of quantitative and qualitative observer rating scales suggested in the literature. All of them center around three key criteria:
1) Was the capability or behavior performed?
2) How well was capability or behavior performed?
3) How often were important behaviors observed?

For example, observers may look for the frequency of behaviors associated with good situational awareness, such as asking for additional information. To assess decision-making abilities, they may rate how well a team adhered to a prescribed decision-making process, or subjectively evaluate the outcome of the decision-making process.

Relevant References
Drills & Exercises Evaluation Guidance Manual. (2010, February). California Department of Fish and Game, Office of Spill Prevention and Response.

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

Cantu, C. J. (2007). *Evaluating team effectiveness: examination of the team assessment tool*. University of North Texas.

Healey, A. N., Undre, S., & Vincent, C. A. (2004). Developing observational measures of performance in surgical teams. *Quality and Safety in Health Care, 13*, i33–i40. http://doi.org/10.1136/qshc.2004.009936

Ott, G. L., Haynes, D. C., & Pond, R. G. (1999). Considering "Best Response" Capabilities in Area Contingency Plans. In *1999 International Oil Spill Conference* (pp. 7–13). American Petroleum Institute.

Rosen, M., Weaver, S., Lazzara, E., Salas, E., Wu, T., Silvestri, S., … King, H. (2010). Tools for evaluating team performance in simulation-based training. *Journal of Emergencies, Trauma, and Shock*, *3*(4), 353. http://doi.org/10.4103/0974-2700.70746

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). Measuring Situation Awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics*, *39*(3), 490–500. http://doi.org/10.1016/j.ergon.2008.10.010

10.1.2   Probing/Question Asking

Description
This evaluation technique describes data collection methods that evaluators can use to help rate performance. There are two primary types of probing/question asking techniques to use in an exercise:

1. **Freeze Probe** – stop exercise and ask questions about the situation to assess understanding
2. **Real-time Probe** – within exercise, ask questions about situation to assess understanding

These two techniques allow evaluators to investigate unspoken knowledge and reasoning in order to more deeply assess competency. For example, in a freeze probe, an evaluator may see a team of exercise participants make a large decision without much discussion. Curious about why and how the participants made that decision, the evaluator may ask about their thought process and understanding of the situation. In a real-time probe aimed at assessing participants' awareness of the situation, the team might receive a phone call from an authority figure away from the site, asking them to summarize the situation.

While the probing techniques are qualitative in nature, they can be used to inform quantitative rating scales as well as a comprehensive qualitative evaluation. For example, after the evaluator asks about the participants thought process and understanding of the situation, the evaluator may record that the decision making process used was "5-Excellent" on the scoring worksheet. A collection of these numbers may be used provide a summative performance score. Any notes taken may also be used to provide a comprehensive qualitative written explanation of exercise performance.

A question guide may be produced in advance or evaluators can rely on their expertise and training when interacting with exercise participants. Regardless, evaluators should remain objective when asking questions so as not to unduly influence exercise play. The primary objective of probing is to clarify behaviors and actions so that a more accurate and informative evaluation can be conducted.

Relevant References

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). Measuring Situation Awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics*, *39*(3), 490–500. http://doi.org/10.1016/j.ergon.2008.10.010

10.1.3   Proxies and Indicators

Description

This evaluation technique utilizes special measures as proxies for assessing performance. It is not always possible to evaluate the exact capability for various reasons, such as the difficulty in measuring a capability, or the artificiality of the exercise. Instead, other measures may be used as proxies for the capability. There are two approaches:

1. Measure performance based on process characteristics (e.g., how good was their decision-making process?)
2. Measure performance based on outcomes (e.g., how good was the outcome of the decision?)

This evaluation technique requires the identification of specific measures that give insight into the performance of a capability. For example, decision-making capabilities could be assessed by evaluating the process of decision-making and the outcome of the decision. As a second example, it may not be possible to evaluate the success of oil spill clean-up because it is not safe to spill oil in an exercise. Instead, an evaluator may evaluate their processes for personnel management and safety, along with the ability to clean up and separate a non-toxic oil substitute.

Relevant References

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). Measuring Situation Awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics*, *39*(3), 490–500. http://doi.org/10.1016/j.ergon.2008.10.010

10.1.4   Self-Rating

Description

This evaluation technique asks participants to rate their own performance in an exercise. Either during or after the exercise, participants complete a pre-developed assessment (paper or electronic). The assessment form can be customized to meet the evaluation needs of exercise designers and include both qualitative as well as quantitative questions.

This form of assessment is easy, but is also highly subjective. However, it can still provide valuable feedback and information to support evaluation, such as measuring learning and enabling participants to identify strengths and areas for improvement. The former can support learning about exercise design and development. The latter can add data to support conclusions in the summative exercise evaluation.

Relevant References

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). Measuring Situation Awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics*, *39*(3), 490–500. http://doi.org/10.1016/j.ergon.2008.10.010

Savoia, E., Testa, M. A., Biddinger, P. D., Cadigan, R. O., Koh, H., Campbell, P., & Stoto, M. A. (2009). Assessing public health capabilities during emergency preparedness tabletop exercises: reliability and validity of a measurement tool. *Public Health Reports*, *124*, 138–148.

10.1.5   Document Review

Description

This evaluation technique asks exercise participants to provide to evaluators the documents used in the exercise. Documents can be used as primary or secondary data to help evaluators assess the performance of capabilities. The goal of this technique is to gather and analyze supplemental data to help evaluate a capability. For example, a resource status report that was provided to Unified Command may be used by an evaluator to help assess the Manage and Account for Resources capability. The evaluator may find that not all requested resources were included in the status report, which provides evidence that the capability was performed, but perhaps not well.

10.1.6   Plan Evaluation

Description

This evaluation technique evaluates the performance of response plans created or utilized in an exercise. The goal of this technique is to highlight aspects of the plan that did and did not work well during the exercise in order to understand how the performance of capabilities was impacted. Plans may be created prior to an exercise or during the exercise in response to an operational need. There are two evaluation approaches:

1. Evaluate the plan for completeness based on whether it covered all contingencies experienced in the exercise.
2. Evaluate the plan for flexibility based on whether it enabled successful response even when contingencies were not specifically planned for.

Carrying out this type of evaluation requires an evaluator to understand both the plan and the actions or processes executed in the exercise. This is easiest to accomplish when an evaluator familiarizes himself or herself with the plan before the exercise, then notes when actions diverge from the plan (and whether there was good reason to do so). These notes should indicate whether unplanned-for contingencies arose during the exercise, and whether the plan enabled flexible response when such contingencies did arise.

Relevant References
Franks, N., Knutson, S., Parker, H. A., & LeJeune, F. (2011). Plan-holder exercises – How to recapture a lost opportunity for area plan improvement. In *2011 International Oil Spill Conference*. American Petroleum Institute.

National Preparedness for Response Exercise Guidelines (DRAFT). (2015, March). USCG, EPA & BSEE.

National Preparedness for Response Exercise Program Guidelines. (2002, August). USCG, EPA, & DOI.

10.1.7 Post-Hoc Modeling

Description
This evaluation technique generates computational models to assess the effectiveness of a simulated response after the exercise has been completed. Models help to assess in greater depth the relative success of the simulated response. There are two ways in which modeling may be used:

1. Model critical aspects of the response and use the model to analyze key statistical measures to identify critical drivers of problems or successes (For example, use systems dynamics to model the decision processes used in the response, then run the model using different available information to determine whether better information collection would enable better decisions. As a second example, use social network analysis to assess coordination and collaboration.)
2. Model the exercise outcomes based on the actions taken and decision made (e.g., conduct an impact analysis to assess the outcomes of all decisions in the exercise)

Modeling is an additional assessment tool that is an effective technique for generating learning about exercises and response operations. The learning can then be used to inform assessment approaches of operator performance in an exercise. For example, these tools may be used after an exercise to identify specific areas for improvement (e.g., how to coordinate better) or understand the impact of certain decisions on outcomes. (e.g., how much did decision 'X' impact outcome 'Y'?) However, this approach is time consuming as it requires experts to build a model for the exercise, collect appropriate data, and analyze the results.

Relevant References
Abrahamsson, M., Hassel, H., & Tehler, H. (2010). Towards a System-Oriented Framework for Analysing and Evaluating Emergency Response. *Journal of Contingencies and Crisis Management*, *18*(1).

Healey, A. N., Undre, S., & Vincent, C. A. (2004). Developing observational measures of performance in surgical teams. *Quality and Safety in Health Care*, *13*, i33–i40. http://doi.org/10.1136/qshc.2004.009936

Iakovou, E., Douligeris, C., & Korde, A. (1994). A Synthesis of Decision Models for Analysis, Assessment, and Contingency Planning for Oil Spill Incidents. *Omega, International Journal of Management Science*, *22*(5), 457–470.

Kapucu, N., & Demiroz, F. (2011). Measuring Performance for Collaborative Public Management Using Network Analysis Methods and Tools. *Public Performance & Management Review*, *34*(4), 549–579. http://doi.org/10.2753/PMR1530-9576340406

Su, Y. S. (2011). Application of Social Network Analysis Methods to Quantitatively Assess Exercise Coordination. *Homeland Security Affairs*, *7*.

10.1.8   Participant Hotwash

Description

This facilitated evaluation technique prompts exercise participants to discuss and clarify information on the response, such as why a task was not completed. In addition, it provides an opportunity for exercise participants to identify initial key strengths and areas for improvement useful for future exercises or responses. Exercise evaluators may then use this information to inform their evaluation of the exercise and the demonstrated capabilities. This is conducted immediately after the exercise ends.

It is useful to structure the hotwash around key issues. With this framework, we recommend structuring the hotwash around the key capabilities evaluated in the exercise.

Relevant References
Drills & Exercises Evaluation Guidance Manual. (2010, February). California Department of Fish and Game, Office of Spill Prevention and Response.

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

National Preparedness for Response Exercise Guidelines (DRAFT). (2015, March). USCG, EPA & BSEE.

National Preparedness for Response Exercise Program Guidelines. (2002, August). USCG, EPA, & DOI.

10.1.9   Controller/Evaluator Debriefing

Description

This facilitated evaluation technique brings together evaluators and other non-playing personnel such as exercise controllers, facilitators, and simulators, to identify and discuss the main strengths and weakness of the response. Exercise participants are not included. The goal is to identify the root causes of issues that will then be used to inform the evaluation of the exercise and the demonstrated capabilities. For example, a controller may have prematurely introduced a change in the situation (e.g., weather update) that

confused exercise participants or resulted in unanticipated response actions. Knowing this allows exercise evaluators to understand whether exercise participants' poor performance was due to their lack of ability or an artificiality of the exercise. This is conducted as soon as possible after the exercise hotwash (e.g., same or next day), but after the participant hotwash.

Relevant References

Drills & Exercises Evaluation Guidance Manual. (2010, February). California Department of Fish and Game, Office of Spill Prevention and Response.

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.

10.1.10 After Action Review

Description

This evaluation technique is designed to qualitatively assess the entire simulated response. Based on the data collected, the assessment of capabilities using the defined measures, and evaluators' expertise, the review process asks evaluators to identify and describe key strengths and areas for improvement, including recommendations. These strengths and areas for improvement are then discussed to understand the root cause of the issue.

The findings are incorporated into a more formal after action report that includes a description of the exercise or response scenario, a discussion of how each capability was performed, a summary of findings and recommendations, and a corrective action plan. The corrective action plan may serve as an agreement between all parties on who is taking responsibility for what corrective actions and by when. This process may be completed immediately following the exercise and take several weeks or months to completed depending on the complexity of the exercise.

Relevant References

Drills & Exercises Evaluation Guidance Manual. (2010, February). California Department of Fish and Game, Office of Spill Prevention and Response.

Homeland Security Exercise and Evaluation Program. (2013, April). U.S. Department of Homeland Security.